Classical Mechanics Lecture Notes

Waseem Bari

December 20, 2017

Abstract

A breif description of the topics taught in Units III and IV shall be presented here. For the detailed description, please consult the class otes as well as the prescribed books for the course.

1 Canonical Transformations

It is already known that Lagrange's equations are invariant under arbitrary *point transformations* i.e., the equations keep their form if we replace q_k by a new set of coordinates Q_k connected with the formaer by relation of the type

$$Q_k = f_k(q_1, q_2, q_3, ..., q_f) \tag{1}$$

The associated P_k are then given by

$$\begin{aligned} P_k &= \frac{\partial L}{\partial \dot{Q}_k} \\ &= \sum_i \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \dot{Q}_k} \\ &= \sum_i p_i a_{ik} \end{aligned} \tag{2}$$

This means that the new momenta are linear functions of p_i whose coefficients are functions of q_k. Transformations of this type are raferred to as the *point transformations*.
Now, we would try to see the scenario under more general transformations i.e.,

$$\begin{aligned} Q_k &= f_k(q, p) \\ P_k &= g_k(q, p) \end{aligned} \tag{3}$$

where f_k and g_k are arbitrary functions of two sets of variables q_k and p_k.
Let us represent the new Hamiltonian as $\bar{H}(Q, P)$.
So we should have in the two sets of coordinates

$$\begin{aligned} H(q, p) &= \sum_k p_k \dot{q}_k - L(q_k, \dot{q}_k) \\ \bar{H}(Q, P) &= \sum_k P_k \dot{Q}_k - \bar{L}(Q_k, \dot{Q}_k) \end{aligned} \tag{4}$$

The above equations reveal that we have two ways of descibing a physical system i.e.

$$L(q_k, \dot{q}_k, t) \text{ and } \bar{L}(Q_k, \dot{Q}_k, t)$$

If the two refer to the same physical system, then

$$\bar{L}(Q_k, \dot{Q}_k, t) = L(q_k, \dot{q}_k, t) - \frac{d}{dt}F(q, Q, t) \tag{5}$$

1.1 The Generating Fnnctions

Let us recall that Hamilton's equations for the original set q_k, p_k of phase space coordinates can be derived on the basis of Hamilton's variational principle

$$\delta \int_{t_1}^{t_2} L\ dt = \delta \int_{t_1}^{t_2} \left(\sum_k p_k dq_k - H \right) dt = 0 \tag{6}$$

The variations δq_k and δp_k are all independent and they are subject to the boundary condition

$$\delta q_k(t = t_0) = \delta q_k(t = t_1) = 0 \tag{7}$$

Now, if the Hamilton's equations are to hold also for new set Q_k, P_k of phase space coordinates, these must follow the Hamilton's principle.
We must, thus, have simultaneously

$$\delta \int_{t=t_0}^{t=t_1} \left(\sum_k P_k dQ_k - \bar{H} \right) dt = 0 \tag{8}$$

Here also, the variations δQ_k and δP_k are taken to be independent and are subjected to the boundary condition

$$\delta Q_k(t = t_0) = \delta Q_k(t = t_1) = 0 \tag{9}$$

Now, in principle, the two formulations are compatible if the integrands

$$\sum_k P_k dQ_k - \bar{H}$$

$$\sum_k p_k dq_k - H$$

differ by the total derivative dF of a function

$$F(q_k, Q_k, t)$$

so that

$$\delta \int_{t_0}^{t_1} dF \ = \ F(q_k(t_1), Q_k(t_1), t_1) - F(q_k(t_0), Q_k(t_0), t_0) \tag{10}$$

By virtue of the boundary conditions, we have

$$\delta \int dF = 0 \tag{11}$$

Thus, the very first condition for the transformation of the phase-space coordinates (q_k, p_k) to a new set of phase-space coordinates (Q_k, P_k) is that there should exist a function $F(q_k, Q_k)$ such that

$$\sum_k p_k \, dq_k - H dt = \sum_k P_k \, dQ_k - \bar{H} dt + dF \tag{12}$$

Te function $F(q_k, Q_k, t)$ is raferred toa s the *Generating function*.
The total time derivative of the funtion F is

$$dF = \sum_k \frac{\partial F}{\partial q_k} dq_k + \sum_k \frac{\partial F}{\partial Q_k} dq_k + \frac{\partial F}{\partial t} dt \tag{13}$$

From Equation (1.12), we may write

$$dF = \sum_k p_k \, dq_k - \sum_k P_k \, dQ_k + (\bar{H} - H) dt \tag{14}$$

Thus comparing Equations (1.13) and (1.14), we obtain the following conditions

$$\begin{aligned} p_k &= \tfrac{\partial F}{\partial q_k} \\ P_k &= -\tfrac{\partial F}{\partial Q_k} \\ \bar{H} &= H + \tfrac{\partial F}{\partial t} \end{aligned} \tag{15}$$

The first two of the above equations give us the old and new momenta in terms of the derivatives of the generating functions whereas the last of the equations gives the transformed Hamiltonian.

Remark: *We made a particular choice to make F a function of q_k and Q_k. Given the symmetry between coordinates and canonical momenta, it is likely that we could equally well write F as a function of (q_k, P_k), (p_k, P_k) or (p_k, Q_k). These different generating functions are simply different ways to generate the same canonical transformation*

$$(q_k, p_k) \longrightarrow (Q_k, P_k)$$

The conjugate pairs remain the same regardless of how the transformation is generated. For reference, we summarize here the four kinds of generating functions and the transformation equations that are derived from them.

$$F_1 : \quad F_1(q_k, Q_k) : \quad p_k = \frac{\partial F_1}{\partial q_k} \ , \ P_k = -\frac{\partial F_1}{\partial Q_k} \tag{16}$$

$$F_2 : \quad F_2(q_k, P_k) : \quad p_k = \frac{\partial F_2}{\partial q_k} \ , \ Q_k = \frac{\partial F_2}{\partial Q_k} \tag{17}$$

$$F_3: \quad F_3(p_k, Q_k): \quad q_k = -\frac{\partial F_3}{\partial p_k} \; , \; P_k = -\frac{\partial F_3}{\partial Q_k} \tag{18}$$

$$F_4: \quad F_4(p_k, P_k): \quad q_k = \frac{\partial F_4}{\partial p_k} \; , \; Q_k = -\frac{\partial F_4}{\partial P_k} \tag{19}$$

1.2 Conditions for Canonical Transformations – I

Most often, it is convenient to test whether a transformation is canonical without attempting to find the actual generating function. In this and the next section, we shall try to look for these direct tests

For simplicity, we assume that the transformatio does not depend explicitly on time. This means that

$$\bar{H} = H \tag{20}$$

Thus, if we have

$$Q_k = Q_k(q_k, p_k)$$

then the time derivative of the new coordinate cna be expressed as

$$\dot{Q}_k = \frac{dQ_k}{dt} \tag{21}$$

which may be written as

$$\dot{Q}_k = \sum_k \left(\frac{\partial Q_k}{\partial q_k} \dot{q}_k + \frac{\partial Q_k}{\partial p_k} \dot{p}_k \right) \tag{22}$$

Now, recall the Hamilton's Equations of motion

$$\begin{aligned} \dot{q}_k &= \tfrac{\partial H}{\partial p_k} \\ \dot{p}_k &= -\tfrac{\partial H}{\partial q_k} \end{aligned} \tag{23}$$

Using these equations, we have from (3.3)

$$\dot{Q}_k = \sum_k \left(\frac{\partial Q_k}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial Q_k}{\partial p_k} \frac{\partial H}{\partial q_k} \right) \tag{24}$$

Now, if we consider the transformation to be canonical, then

$$\dot{Q}_k = \frac{\partial \bar{H}}{\partial P_k} = \frac{\partial H}{\partial P_k} \tag{25}$$

which implies that

$$\dot{Q}_k - \frac{\partial H}{\partial P_k} = 0 \tag{26}$$

However,

$$H = H(q_k, p_k)$$

we may write

$$\frac{\partial H}{\partial P_k} = \sum_k \left(\frac{\partial H}{\partial q_k}\frac{\partial q_k}{\partial P_k} + \frac{\partial H}{\partial p_k}\frac{\partial p_k}{\partial P_k} \right) \tag{27}$$

Using Equations (3.5) and (3.8) in (3.7), we have

$$\sum_k \left(\frac{\partial Q_k}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial Q_k}{\partial p_k}\frac{\partial H}{\partial q_k} \right) - \sum_k \left(\frac{\partial H}{\partial q_k}\frac{\partial q_k}{\partial P_k} + \frac{\partial H}{\partial p_k}\frac{\partial p_k}{\partial P_k} \right) = 0 \tag{28}$$

which on simplification gives

$$\sum_k \left(\frac{\partial Q_k}{\partial q_k} - \frac{\partial p_k}{\partial P_k} \right)\frac{\partial H}{\partial p_k} - \sum_k \left(\frac{\partial q_k}{\partial P_k} + \frac{\partial Q_k}{\partial p_k} \right)\frac{\partial H}{\partial q_k} = 0 \tag{29}$$

The above equation would be satisfired if the coefficients of $\frac{\partial H}{\partial p_k}$ and $\frac{\partial H}{\partial q_k}$ vanish separately. This would lead to

$$\begin{aligned} \frac{\partial Q_k}{\partial q_k} - \frac{\partial p_k}{\partial P_k} &= 0 \\ \frac{\partial q_k}{\partial P_k} + \frac{\partial Q}{\partial p_k} &= 0 \end{aligned} \tag{30}$$

The above equations lead to the following conditions for a transformation $(q_k, p_k) \longrightarrow (Q_k, P_k)$ to be canonical

$$\begin{aligned} \frac{\partial Q_k}{\partial q_k} &= \frac{\partial p_k}{\partial P_k} \\ \frac{\partial q_k}{\partial P_k} &= -\frac{\partial Q_k}{\partial p_k} \end{aligned} \tag{31}$$

This gives one set of conditions that can be used to test whether a transformation is canonical or not. The second set of conditions is derived in the next section.

1.3 Conditions for Canonical Transformations – II

The second set of conditions is obtained by starting with the assumption that

$$P_k = P_k(q_k, p_k)$$

so that one may write the time derivative as

$$\dot{P}_k = \sum_k \left(\frac{\partial P_k}{\partial q_k}\dot{q}_k + \frac{\partial P_k}{\partial p_k}\dot{p}_k \right) \tag{32}$$

Using the Hamilton's equations of motion, the above can be rewritten as

$$\dot{P}_k = \sum_k \left(\frac{\partial P_k}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial P_k}{\partial p_k}\frac{\partial H}{\partial q_k} \right) \tag{33}$$

For the transformation to be canonical, we should have

$$\dot{P}_k = -\frac{\partial \bar{H}}{\partial Q_k} = -\frac{\partial H}{\partial Q_k} \tag{34}$$

which gives

$$\dot{P}_k + \frac{\partial H}{\partial Q_k} = 0 \tag{35}$$

However since

$$H = H(q_k, p_k)$$

we have

$$\frac{\partial H}{\partial Q_k} = \sum_k \left(\frac{\partial H}{\partial q_k}\frac{\partial q_k}{\partial Q_k} + \frac{\partial H}{\partial p_k}\frac{\partial p_k}{\partial Q_k} \right) \tag{36}$$

Using Equations (4.2) and (4.5) in Equation (4.4), we have

$$\sum_k \left(\frac{\partial P_k}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial P_k}{\partial p_k}\frac{\partial H}{\partial q_k} \right) + \sum_k \left(\frac{\partial H}{\partial q_k}\frac{\partial q_k}{\partial Q_k} + \frac{\partial H}{\partial p_k}\frac{\partial p_k}{\partial Q_k} \right) = 0 \tag{37}$$

which upon simplification gives

$$\sum_k \left(\frac{\partial P_k}{\partial q_k} + \frac{\partial p_k}{\partial Q_k} \right) - \sum_k \left(\frac{\partial P_k}{\partial p_k} - \frac{\partial q_k}{\partial Q_k} \right) \frac{\partial H}{\partial q_k} = 0 \tag{38}$$

which gives the following conditions

$$\begin{aligned} \frac{\partial P_k}{\partial q_k} &= \frac{\partial p_k}{\partial Q_k} \\ \frac{\partial P_k}{\partial p_k} &= -\frac{\partial q_k}{\partial Q_k} \end{aligned} \tag{39}$$

This is another set of conditions which must be met of a transformation is canonical.

1.4 The Harmonic Ocillator

Example: The Harmonic Oscillator
This example provides the complete description of the Harmonic Oscillator in terms of the generating functions $F = F_1(q, Q) = \frac{1}{2}\omega q^2 \cot 2\pi Q_k$.
Solution: The Hamiltonian for a Harmonic Oscillator is written as

$$H = \frac{1}{2}\left(p^2 + \omega^2 q^2\right) \tag{40}$$

where

$$q = \sqrt{m}x \tag{41}$$

which implies that

$$p = \sqrt{m}\dot{x} \tag{42}$$

Further,

$$\omega^2 = \frac{k}{m} \tag{43}$$

Now, let us discuss the different forms of the generating functions separately as follows We take the case when $F = F_1(q, Q)$ with the following form

$$F = F(q, Q) = \frac{1}{2}\omega q^2 \cot 2\pi Q_k \tag{44}$$

Now, let us recall the conditions (1.15) for transformation to be canonical

$$\begin{aligned} p_k &= \tfrac{\partial F}{\partial q_k} \\ P_k &= -\tfrac{\partial F}{\partial Q_k} \\ \bar{H} &= H + \tfrac{\partial F}{\partial t} \end{aligned} \tag{45}$$

Using first two of the above equations, we get in the present case,

$$p_k = \frac{\partial F}{\partial q_k} = \frac{\partial}{\partial q_k}\left[\frac{1}{2}\omega q^2 \cot 2\pi Q_k\right] \tag{46}$$

which simplifies to

$$p_k = \omega q_k \cot 2\pi Q_k \tag{47}$$

Using second of the conditions

$$P_k = -\frac{\partial F}{\partial Q_k} = \frac{\partial}{\partial Q_k}\left[\frac{1}{2}\omega q^2 \cot 2\pi Q_k\right] \tag{48}$$

which simplifies to

$$P_k = \pi\omega q_k^2 \text{ cosec}^2 2\pi Q_k \tag{49}$$

Now, to actually visulaize the transformation, we need to write (q, p) in terms of (Q, P). For this purpose, let us initially invert Equation (1.49) so that

$$q_k = \sqrt{\frac{P_k}{\pi\omega}} \sin 2\pi Q_k \tag{50}$$

Using the above in Equation (1.47), we have

$$p_k = \omega\sqrt{\frac{P_k}{\pi\omega}} \sin 2\pi Q_k \cot 2\pi Q_k \tag{51}$$

which on simplification leads to

$$p_k = \sqrt{\frac{\omega P_k}{\pi}} \cos 2\pi Q_k \tag{52}$$

Thus, we have the following two transformation relations

$$q_k = \sqrt{\frac{P_k}{\pi\omega}} \sin 2\pi Q_k \tag{53}$$

and

$$p_k = \sqrt{\frac{\omega P_k}{\pi}} \cos 2\pi Q_k \tag{54}$$

Now, let us focus on the transformed Hamiltonian

$$\bar{H}(Q,P) = H(q,p) + \frac{\partial F}{\partial t} = \frac{1}{2}\left(p^2 + \omega^2 q^2\right) + \frac{\partial F}{\partial t} \tag{55}$$

Substitute the transformation relations (1.53) and (1.54) in the above, we have the transformed Hamiltonian of the form

$$\bar{H}(Q,P) = \frac{1}{2}\left[\frac{\omega P_k}{\pi}\cos^2 2\pi Q_k + \omega^2 \frac{P_k}{\pi\omega}\sin^2 2\pi Q_k\right] \tag{56}$$

which can easily be simplified to

$$\bar{H}(Q,P) = \frac{\omega}{\pi} P_k \tag{57}$$

Let us now determine the Equations of motion for Q_k and P_k using the Hamilton's Equations

$$\begin{aligned} \dot{Q}_k &= \tfrac{\partial \bar{H}}{\partial P_k} \\ \dot{P}_k &= -\tfrac{\partial \bar{H}}{\partial Q_k} \end{aligned} \tag{58}$$

so that using Equation (1.57)Integration of the above gives us

$$\begin{aligned} Q_k &= \tfrac{\omega t}{2\pi} \\ P_k &= \text{Constant} = P_0 \end{aligned} \tag{59}$$

This means that we have found Q_k and P_k as explicit functions of time.
In light of the above, the Equations (1.53) and (1.54) can be written as

$$q_k = \sqrt{\frac{P_0}{\pi\omega}} \sin \omega t \tag{60}$$

and

$$p_k = \sqrt{\frac{\omega P_0}{\pi}} \cos \omega t \tag{61}$$

We have, thus, transformed from simple position and monetum to phase (Q_k) and energy (P_k) of the oscillatory motion.

Remark: *There can be other forms of the generating funtions (G.F.) like* $F = F_2(q, P)$*,* $F = F_3(p, Q)$ *and* $F = F_4(p, P)$*. The students should consult the class notes to derive the following transformation relations in the various cases of the generating functions.*

If we consider $F = F_2(q, P)$ with the following form

$$F = F_2(q, P) = F_1(q, Q) + QP \tag{62}$$

We have the fundamental Poisson Brackets

$$\begin{aligned} [q_k, q_l] &= 0 \\ [p_k, p_l] &= 0 \\ [p_k, q_l] &= -\delta_{kl} \\ [q_k, p_l] &= \delta_{kl} \end{aligned} \tag{63}$$

2 Poisson Brackets

Let us consider the transformation

$$(q_k, p_k) \longrightarrow (Q_k, P_k)$$

such that

$$\begin{aligned} Q_k &= q_k + \delta q_k \\ P_k &= p_k + \delta p_k \end{aligned} \tag{64}$$

i.e., the transfromed set of coordinates differs from that of the original ones by infinitesimals. Such transformations are known as the *Infinitesimal Transformations*
To have a better understanding of thses transformations, let us consider the following generating function

$$F_2(q_k, P_k) = \sum_k q_k P_k \tag{65}$$

Now, if we apply the canonical transformations, (Refer to Equation (2.12)), we have

$$p_k = \frac{\partial F_2}{\partial q_k} = P_k \tag{66}$$

$$Q_k = \frac{\partial F_2}{\partial P_k} = q_k \tag{67}$$

and

$$\bar{H} = H \quad (F_2 \text{ is independent of } t) \tag{68}$$

This means that the generating function $\sum_k q_k P_k$ generates the transformations of the form

$$\begin{aligned} Q_k &= q_k \\ p_k &= P_k \end{aligned} \quad (69)$$

Transformations of this sort are known as the *Identical Transformations*.
Now, a generating function which gives rise to an infinitesimal transformation i.e., an infinitesimal change in the variables, can be written as

$$F_2(q_k, P_k) = \sum_k q_k P_k + \epsilon G(q_k, P_k) \quad (70)$$

where $G(q_k, P_k)$ is an arbitrary function and ϵ is infinitesimally small.
If the generating function is of the above form, the canonical transformation leads to

$$Q_k = \frac{\partial F_2}{\partial P_k} = q_k + \epsilon \frac{\partial G}{\partial P_k} \quad (71)$$

and

$$p_k = \frac{\partial F_2}{\partial q_k} = P_k + \epsilon \frac{\partial G}{\partial q_k} \quad (72)$$

The above two equations thus lead to

$$Q_k - q_k = \delta q_k = \epsilon \frac{\partial G}{\partial P_k} \quad (73)$$

and

$$P_k - p_k = \delta p_k = -\epsilon \frac{\partial G}{\partial q_k} \quad (74)$$

Since $P_k - p_k$ is infinitesimally small, it is possible to replace P_k by p_k in the derivative as well as in $G(q_k, P_k)$, so that the above two equations can be rewritten as

$$\delta q_k = \epsilon \frac{\partial G}{\partial p_k} \quad (75)$$

and

$$\delta p_k = -\epsilon \frac{\partial G}{\partial q_k} \quad (76)$$

Now, if we consider any function $F(q_k, p_k)$, the change in F with the changes δq_k and δp_k in q_k and p_k respectively can be expressed as

$$\delta F = \sum_k \left[\frac{\partial F}{\partial q_k} \delta q_k + \frac{\partial F}{\partial p_k} \delta p_k \right] \quad (77)$$

Using the Equations (5.12) and (5.13) in the above equation, we get

$$\delta F = \epsilon \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} \right] \quad (78)$$

The quantity in the brackets above is known as the *Poisson Bracket* of two functions or dynamic variables $F(q_k, p_k)$ and $G(q_k, p_k)$ and is denoted by

$$[F, G]_{q_k,p_k}$$

so that

$$[F, G]_{q_k,p_k} = \sum_k \left[\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k} \right] \tag{79}$$

For a single degree of freedom the Poission Bracket is written as

$$[F, G] = \frac{\partial F}{\partial q}\frac{\partial G}{\partial p} - \frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q} \tag{80}$$

2.1 Significance of defining a Poisson Bracket

We still would like to know about the significance of defining such a quantity. For the purpose, let us start with the basic fact that the generating function is a function of q_k, p_k and t

$$F = F(q_k, p_k, t) \tag{81}$$

so that the total time derivative of the generating function is written as

$$\frac{dF}{dt} = \sum_k \left[\frac{\partial F}{\partial q_k}\dot{q}_k + \frac{\partial F}{\partial p_k}\dot{p}_k \right] + \frac{\partial F}{\partial t} \tag{82}$$

Let us recall the Hamilton's Equations of motion

$$\begin{aligned} \dot{q}_k &= \tfrac{\partial H}{\partial p_k} \\ \dot{p}_k &= -\tfrac{\partial H}{\partial q_k} \end{aligned} \tag{83}$$

Using the above equations in Equation (5.19) gives us

$$\frac{dF}{dt} = \sum_k \left[\frac{\partial F}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial H}{\partial q_k} \right] + \frac{\partial F}{\partial t} \tag{84}$$

which in light of the definition of the Poisson Brackets can bew written as

$$\frac{dF}{dt} = [F, H] + \frac{\partial F}{\partial t} \tag{85}$$

The above equation reveasl that F is a constant of motion if

$$[F, H] + \frac{\partial F}{\partial t} = 0 \tag{86}$$

And if F does not have an explicit dependence on time (in that case $\frac{\partial F}{\partial t} = 0$), then the condition for F to be constant of motion is

$$[F, H] = 0 \tag{87}$$

Thus, we conclude that, if a function F does not depend on time explicitly and is a constant of motion, then its Poisson Bracket with the Hamiltonian vanishes.
In other words, a function whose Poisson Bracket with the Hamiltonian vanishes is a constant of motion.

2.2 Equations of Motion in terms of Poisson Brackets

We already knwo the Hamilton'e Equations of motion are

$$\begin{aligned} \dot{q}_k &= \frac{\partial H}{\partial p_k} \\ \dot{p}_k &= -\frac{\partial H}{\partial q_k} \\ -\frac{\partial L}{\partial t} &= \frac{\partial H}{\partial t} \end{aligned} \tag{88}$$

Now recall the equation

$$\frac{dF}{dt} = [F, H] + \frac{\partial F}{\partial t} \tag{89}$$

In the above Equation, if we put

$$F = q_k$$

, then we have

$$\dot{q}_k = [q_k, H] \tag{90}$$

And if we substitute

$$F = p_k$$

, then we have

$$\dot{p}_k = [p_k, H] \tag{91}$$

And finally if

$$F = H$$

, then we have

$$\dot{H} = \frac{\partial H}{\partial t} \tag{92}$$

Thus, with the above discussion, we may write the Hamilton's equations of motion in Poisson Bracket form as under

$$\begin{aligned} \dot{q}_k &= [q_k, H] \\ \dot{p}_k &= [p_k, H] \end{aligned} \tag{93}$$

2.3 Fundamental Poisson Brackets

Let us recall the general definition of the Poisson Brackets

$$[F, G]_{q_k, p_k} = \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} \right] \tag{94}$$

Let us try to see the scenario in some special cases as follows

(a) If $G = q_l$, then we have

$$[F, q_l] = \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial q_l}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial q_l}{\partial q_k} \right] \tag{95}$$

Which simplifies to

$$[F, q_l] = -\sum_k \frac{\partial F}{\partial p_k} \delta_{lk} \tag{96}$$

so that

$$[F, q_l] = -\frac{\partial F}{\partial p_l} \tag{97}$$

In the above result, if we put $F = q_k$, then we get

$$[q_k, q_l] = -\frac{\partial q_k}{\partial p_l} \tag{98}$$

which reduces to

$$[q_k, q_l] = 0 \tag{99}$$

And if $F = p_k$, we have

$$[p_k, q_l] = -\frac{\partial p_k}{\partial p_l} \tag{100}$$

which may be written as

$$[p_k, q_l] = -\delta_{kl} \tag{101}$$

(b) If $G = p_l$, then we have

$$[F, p_l] = \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial p_l}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial p_l}{\partial q_k} \right] \tag{102}$$

Which simplifies to

$$[F, p_l] = \sum_k \frac{\partial F}{\partial q_k} \delta_{lk} \tag{103}$$

so that

$$[F, p_l] = \frac{\partial F}{\partial p_l} \tag{104}$$

In the above result, if we put $F = p_k$, then we get

$$[p_k, p_l] = \frac{\partial p_k}{\partial p_l} \tag{105}$$

which reduces to

$$[p_k, p_l] = 0 \tag{106}$$

And if $F = q_k$, we have

$$[q_k, p_l] = \frac{\partial q_k}{\partial q_l} \tag{107}$$

which may be written as

$$[q_k, p_l] = \delta_{kl} \tag{108}$$

With the above dicussion in (a) and (b), we have the following identities

$$\begin{aligned} [q_k, q_l] &= 0 \\ [p_k, p_l] &= 0 \\ [p_k, q_l] &= -\delta_{kl} \\ [q_k, p_l] &= \delta_{kl} \end{aligned} \tag{109}$$

The above Poisson Brackets are known as the fundamental Poisson Brackets.

2.4 Invariance of Poisson Brackets w.r.t. Canonical Transformations

Poisson Brackets are invariant under canonical transformations. Let us first have a look at the fundamental Poisson Brackets described by Equation (5.46). Let us recall the general definition of the Poisson Brackets

$$[F, G]_{q_k, p_k} = \sum_k \left[\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k} \right] \tag{110}$$

Consider the Canonical Transformation

$$(q_k, p_k) \longrightarrow (Q_k, P_k)$$

The above transformation implies that

$$\begin{aligned} q_k &= q_k(Q_k, p_k) \\ p_k &= p_k(Q_k, p_k) \end{aligned} \tag{111}$$

Therefore, one can write

$$\frac{\partial G}{\partial q_k} = \sum_k \left[\frac{\partial G}{\partial Q_l}\frac{\partial Q_l}{\partial q_k} + \frac{\partial G}{\partial P_l}\frac{\partial P_l}{\partial q_k}\right] \tag{112}$$

and

$$\frac{\partial G}{\partial p_k} = \sum_k \left[\frac{\partial G}{\partial Q_l}\frac{\partial Q_l}{\partial p_k} + \frac{\partial G}{\partial P_l}\frac{\partial P_l}{\partial p_k}\right] \tag{113}$$

Thus, we can write Equation (5.47) with the above results as

$$[F,G]_{q,p} = \sum_k \sum_l \left[\frac{\partial F}{\partial q_k}\left\{\frac{\partial G}{\partial Q_l}\frac{\partial Q_l}{\partial p_k} + \frac{\partial G}{\partial P_l}\frac{\partial P_l}{\partial p_k}\right\} - \frac{\partial F}{\partial p_k}\left\{\frac{\partial G}{\partial Q_l}\frac{\partial Q_l}{\partial q_k} + \frac{\partial G}{\partial P_l}\frac{\partial P_l}{\partial q_k}\right\}\right] \tag{114}$$

The above can be simplified as

$$[F,G]_{q,p} = \sum_k \sum_l \left[\frac{\partial G}{\partial Q_l}\left\{\frac{\partial F}{\partial q_k}\frac{\partial Q_l}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial Q_l}{\partial q_k}\right\} + \frac{\partial G}{\partial P_l}\left\{\frac{\partial F}{\partial q_k}\frac{\partial P_l}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial P_l}{\partial q_k}\right\}\right] \tag{115}$$

Which can be writtes as

$$[F,G]_{q,p} = \sum_l \left[\frac{\partial G}{\partial Q_l}[F,Q_l]_{q_k,p_k} + \frac{\partial G}{\partial P_l}[F,P_l]_{q_k,p_k}\right] \tag{116}$$

In the above Equation, if we put $F = Q_i$ and $G = F$, we get

$$[Q_i,F]_{q,p} = \sum_l \left[\frac{\partial F}{\partial Q_l}[Q_i,Q_l]_{q_k,p_k} + \frac{\partial F}{\partial P_l}[Q_i,P_l]_{q_k,p_k}\right] \tag{117}$$

Using the fundamental Poisson Brakcet i.e. $[Q_i,Q_l]_{q_k,p_k} = 0$ and $[Q_i,P_l]_{q_k,p_k} = \delta_{kl}$, we can simplify the above equation to

$$[Q_i,F]_{q,p} = \sum_l \left[\frac{\partial F}{\partial P_l}\delta_{il}\right] \tag{118}$$

which reduces to

$$[Q_i,F]_{q,p} = \frac{\partial F}{\partial P_i} \tag{119}$$

In a similar fashion, if we put $F = P_i$ and $G = F$ in Equation (5.53), we can easily reach to the result

$$[P_i,F]_{q,p} = -\frac{\partial F}{\partial Q_i} \tag{120}$$

Using the Equations (5.56) and (5.57) in (5.53), we have

$$[F,G]_{q,p} = \sum_l \left[-\frac{\partial G}{\partial Q_l}\frac{\partial F}{\partial P_i} + \frac{\partial G}{\partial P_l}\frac{\partial F}{\partial Q_i}\right] \tag{121}$$

which gives

$$[F, G]_{q,p} = [F, G]_{Q,P} \tag{122}$$

Thus, we see that the Poisson Brackets are invariant under the Canonical transformations.

2.5 Poisson Brackets and Constants of Motion

The Poisson bracket displays its true power in the search for constants of the motion. A constant of the motion is some function of phase space, independent of time, $F(q_i, p_i)$, whose value is constant for any particle. In other words, $F(q_i, p_i)$ is a constant of the motion if $\frac{dF}{dt} = 0$. Since we specified that F does not depend explicitly in time it follows that $\frac{\partial F}{\partial t} = 0$. This means that $F$is a constant of the motion if and only if $[F, H] = 0$ for all points in phase space. All of the familiar constants of the motion can be checked using this one simple prescription as described below

- **Energy:** First of all we should keep in mind that due to the antisymmetry of the Poisson bracket, we always have

$$[H, H] = 0 \tag{123}$$

Using the following relation

$$\frac{dF}{dt} = [F, H] + \frac{\partial F}{\partial t} \tag{124}$$

we can quickly find that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \tag{125}$$

In those cases where Hamiltonian does not depend on time explicitly $\frac{\partial H}{\partial t} = 0$. Using this in the above equation gives the immediate result that $H(q_i, p_i)$ is a constant of the motion. Thus **Energy is conserved in cases where the Hamiltonian is time-independent.**

- **Linear Momentum:** In a case where the Hamiltonian does not contain a particular coordinate, q_k , explicitly it is said to be cyclic in that coordinate. Applying the the general definition of Poisson bracket i.e.

$$[F, G]_{q,p} = \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} \right] \tag{126}$$

$$[p_k, H]_{q,p} = \sum_k \left[\frac{\partial p_k}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial p_k}{\partial p_k} \frac{\partial H}{\partial q_k} \right] \tag{127}$$

giving us

$$[p_k, H]_{q,p} = \sum_k \left[\frac{\partial H}{\partial q_k} \right] \tag{128}$$

Since, as stated earlier, Hamiltonian does not contain a particular coordinate, q_k , we have

$$[p_k, H]_{q,p} = 0 \tag{129}$$

The above result reveasl that p_k is a constant of motion. Thus, **Momentum is conserved if it is conjugate to a cyclic coordinate.**

- **Angular Momentum:** Consider a particle in three dimension, (x, y, z), subject to a central force potential

$$V(x, y, z) = V(\sqrt{x^2 + y^2 + z^2}) \tag{130}$$

where

$$r = \sqrt{x^2 + y^2 + z^2}$$

is distance from the origin. The Hamiltonian can be written as the sum of the kinetic and potential energy

$$H(p_x, p_y, p_z, x, y, z) = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + T(p_x, p_y, p_z) \tag{131}$$

Angular momentum about the z-axis is defined by the function

$$L_z = xp_y - yp_x$$

Now, linearity allows us to break the Poisson bracket with the Hamiltonian into two smaller pieces

$$[L_z, H] = [L_z, T + V] = [L_z, T] + [L_z, V] \tag{132}$$

Since T depends only on momenta the first bracket on the right has only two non-vanishing terms as described below:

$$\begin{aligned}
[L_z, T] &= \tfrac{\partial L_z}{\partial x}\tfrac{\partial T}{\partial p_x} - \tfrac{\partial L_z}{\partial p_x}\tfrac{\partial T}{\partial x} + \tfrac{\partial L_z}{\partial y}\tfrac{\partial T}{\partial p_y} - \tfrac{\partial L_z}{\partial p_y}\tfrac{\partial T}{\partial y} + \tfrac{\partial L_z}{\partial z}\tfrac{\partial T}{\partial p_z} - \tfrac{\partial L_z}{\partial p_z}\tfrac{\partial T}{\partial z} \\
&= \tfrac{\partial L_z}{\partial x}\tfrac{\partial T}{\partial p_x} + \tfrac{\partial L_z}{\partial y}\tfrac{\partial T}{\partial p_y} \\
&= p_y \tfrac{p_x}{m} - p_x \tfrac{p_y}{m} \\
&= 0
\end{aligned} \tag{133}$$

Other terms, such as $\frac{\partial L_z}{\partial p_x}\frac{\partial T}{\partial x}$ vanish since derivatives of T w.r.t. coordinates, such as x, will always be zero.
Now since V depends only on coordinates so the second bracket in (2.89) has two different non-vanishing terms as described below:

$$
\begin{aligned}
[L_z, V] &= \frac{\partial L_z}{\partial x}\frac{\partial V}{\partial p_x} - \frac{\partial L_z}{\partial p_x}\frac{\partial V}{\partial x} + \frac{\partial L_z}{\partial y}\frac{\partial V}{\partial p_y} - \frac{\partial L_z}{\partial p_y}\frac{\partial V}{\partial y} + \frac{\partial L_z}{\partial z}\frac{\partial V}{\partial p_z} - \frac{\partial L_z}{\partial p_z}\frac{\partial V}{\partial z} \\
&= -\frac{\partial L_z}{\partial p_x}\frac{\partial V}{\partial x} - \frac{\partial L_z}{\partial p_y}\frac{\partial V}{\partial y} \\
&= y\frac{\partial V}{\partial x} - x\frac{\partial V}{\partial y}
\end{aligned} \tag{134}
$$

Using the chain rule, we may write

$$
\begin{aligned}
\frac{\partial V}{\partial x} &= \frac{x}{\sqrt{x^2+y^2+z^2}} V'\left(\sqrt{x^2+y^2+z^2}\right) \\
\frac{\partial V}{\partial y} &= \frac{y}{\sqrt{x^2+y^2+z^2}} V'\left(\sqrt{x^2+y^2+z^2}\right)
\end{aligned} \tag{135}
$$

Thus, using (2.91) gives us

$$
\begin{aligned}
[L_z, V] &= y\frac{\partial V}{\partial x} - x\frac{\partial V}{\partial y} \\
&= y\frac{x}{\sqrt{x^2+y^2+z^2}} V'\left(\sqrt{x^2+y^2+z^2}\right) - x\frac{y}{\sqrt{x^2+y^2+z^2}} V'\left(\sqrt{x^2+y^2+z^2}\right) \\
&= 0
\end{aligned} \tag{136}
$$

Using (2.90) and (2.93) in (2.89) gives us

$$
[L_z, H] = 0 \tag{137}
$$

This means that the z component of the angular momentum is a constant of motion. Using a similar approach, we can show that the x and y components of the angular momentum given by

$$
L_x = yp_z - zp_x \text{ and } L_y = zp_x - xp_z \tag{138}
$$

are also constants of motion i.e.

$$
\begin{aligned}
[L_x, H] &= 0 \\
[L_y, H] &= 0
\end{aligned} \tag{139}
$$

Therefore for a particle moving in a central force potential **all three components of angular momentum are conserved.**

2.6 The Angular Momentum Poisson Brackets Relations

Consider the angular momentum

$$\vec{L} = \vec{r} \times \vec{p}$$

which in componet form reads as

$$L_x = yp_z - zp_x \ , \ L_y = zp_x - xp_z \ , \ L_z = xp_y - yp_x \tag{140}$$

Please consult the class notes to reach to the following relations

$$\begin{aligned} [L_x, p_y] &= p_z \\ [L_y, p_z] &= p_x \\ [L_z, p_x] &= p_y \end{aligned} \tag{141}$$

$$\begin{aligned} [L_x, p_x] &= 0 \\ [L_y, p_y] &= 0 \\ [L_z, p_z] &= 0 \end{aligned} \tag{142}$$

$$\begin{aligned} [L_x, p_z] &= -p_y \\ [L_y, p_x] &= -p_z \\ [L_z, p_y] &= -p_x \end{aligned} \tag{143}$$

semimajor axes $\sqrt{\frac{2\,I}{\omega}}$ and $\sqrt{2\,I\,\omega}$, so the area is $2\,\pi\,I$ as one would expect. I is to first order independent of the variation of ω, so the area of the orbit stays fixed. The ellipse changes shape – if ω is increased, the maximum amplitude of the motion decreases and the maximum momentum increases – but its area is preserved. This is analogous to Liouville's theorem.

Other Examples: see Hand and Finch Section 6.5 for the pendulum example.

2.4.5 The Hamilton-Jacobi Equation

The Hamilton-Jacobi equation makes use of a special canonical transformation to convert the standard Hamiltonian problem of $2M$ first-order ordinary differential equations in $2M$ variables into a single first-order partial differential equation with $M+1$ partial derivatives with respect to the $\{q_k\}$ and time.

The Goal

We propose to do a slightly crazy thing – we want a canonical transformation from some arbitrary set of generalized coordinates $(\vec{q},\vec{p})$ to some new set $(\vec{Q},\vec{P})$ such that all the $\vec{Q}$ and $\vec{P}$ are constant. One way to guarantee this is to require that $\tilde{H}(\vec{Q},\vec{P})=0$. Recalling our equation for the transformation of the Hamiltonian under a canonical transformation, Equation 2.41, we are requiring there be a generating function F such that

$$0=\tilde{H}(\vec{Q},\vec{P})=H(\vec{q},\vec{p})+\frac{\partial F}{\partial t}$$

This is essentially a differential equation for F. Is it possible to find such a function F?

The Formal Solution – the Hamilton-Jacobi Equation

Since the new momenta will be constant, it is sensible to make F a function of the type F_2, $F=S(\vec{q},\vec{P})$. The $\vec{p}$ thus satisfy $p_k=\frac{\partial S}{\partial q_k}$. Our condition on the generating function is thus the partial differential equation, also known as the **Hamilton-Jacobi Equation**,

$$H\left(\vec{q},\frac{\partial S(\vec{q},\vec{P})}{\partial\vec{q}},t\right)+\frac{\partial S(\vec{q},\vec{P})}{\partial t}=0 \tag{2.61}$$

S is known as **Hamilton's Principal Function**. Since the $\vec{P}$ are constants, this is a partial differential equation in $M+1$ independent variables $\vec{q}$ and t for the function S. We are in this case choosing not to consider the partial derivatives $\frac{\partial S}{\partial \vec{q}}=\vec{p}$ to be independent of $\vec{q}$.[15] Since we have $M+1$ independent variables, there are $M+1$ constants of integration. One of these is the constant offset of S, which is physically irrelevant because physical quantities depend only on partial derivatives of S. Since a solution S of this equation will generate a transformation that makes the M components of $\vec{P}$ constant, and since S is a function of the $\vec{P}$, the $\vec{P}$ can be taken to be the M constants.[16]

Independent of the above equation, we know that there must be M additional constants to specify the full motion. These are the $\vec{Q}$. The existence of these extra constants is not

[15] As we have discussed many times before, it is our choice whether to impose this "constraint" at the beginning or the end of solving the problem; if we did not impose this constraint at the beginning, we would have to carry along the M constraint equations $\vec{p}=\frac{\partial S}{\partial \vec{q}}$ and apply them at the end.

[16] One could choose an arbitrary information-preserving combination of the $\vec{P}$ to instead be the constants, but clearly the choice we have made is the simplest one.

implied by the Hamilton-Jacobi equation, since it only needs $M + 1$ constants to find a full solution S. The additional M constants exist because of Hamilton's equations, which require $2\,M$ initial conditions for a full solution.

Since the $\vec{P}$ and $\vec{Q}$ are constants, it is conventional to refer to them with the symbols $\vec{\alpha} = \vec{P}$ and $\vec{\beta} = \vec{Q}$. The full solution $(\vec{q}(t), \vec{p}(t))$ to the problem is found by making use of the generating function S and the initial conditions $\vec{q}(0)$ and $\vec{p}(0)$. Recall the generating function partial derivative relations are

$$\vec{p} = \frac{\partial S(\vec{q}, \vec{\alpha}, t)}{\partial \vec{q}} \tag{2.62}$$

$$\vec{\beta} = \frac{\partial S(\vec{q}, \vec{\alpha}, t)}{\partial \vec{\alpha}} \tag{2.63}$$

The constants $\vec{\alpha}$ and $\vec{\beta}$ are found by applying the above equations at $t = 0$:

$$\vec{p}(t=0) = \left.\frac{\partial S(\vec{q}, \vec{\alpha}, t)}{\partial \vec{q}}\right|_{t=0, \vec{q}(t=0), \vec{\alpha}} \tag{2.64}$$

$$\vec{\beta} = \left.\frac{\partial S(\vec{q}, \vec{\alpha}, t)}{\partial \vec{\alpha}}\right|_{t=0, \vec{q}(t=0), \vec{\alpha}} \tag{2.65}$$

The latter two equations let us determine the $\vec{\alpha}$ and $\vec{\beta}$ from $\vec{q}(t = 0)$ and $\vec{p}(t = 0)$, and then the former two equations give us (implicitly, at least) $\vec{q}(t)$ and $\vec{p}(t)$.

We can more directly determine what S is by evaluating its total time derivative:

$$\begin{aligned} \frac{dS}{dt} &= \frac{\partial S}{\partial t} + \sum_k \frac{\partial S}{\partial q_k}\,\dot{q}_k \\ &= -H + \sum_k p_k\,\dot{q}_k \\ &= L \end{aligned}$$

where we arrived at the second line by using one of the generating function partial derivative relations $p_k = \frac{\partial S}{\partial q_k}$ and the Hamilton-Jacobi equation. We thus see that the generating function S is just the indefinite integral of the Lagrangian:

$$S = \int dt\,L$$

This is an interesting result – that the action integral is the generator of the canonical transformation that corresponds to time evolution. It comes back as a defining principle in field theory, quantum mechanics, and quantum field theory, and, to some extent, in the Feynman path integral formulation of quantum mechanics. It is not of particular practical use, though, since to calculate the indefinite integral we must already know the solution $\vec{q}(t), \vec{p}(t)$.

When H is Conserved – Hamilton's Characteristic Function and the Abbreviated Action

Let us consider the common case of H being conserved. This certainly occurs if H has no explicit time dependence, though that is not a necessary condition. Since H is a constant, we know that S can be written in the form

$$S(\vec{q}, \vec{\alpha}, t) = W(\vec{q}, \vec{\alpha}) - E\,t \tag{2.66}$$

where $E = H$ is the time-independent value of H. That the above rewriting is possible is seen by simply calculating $\frac{\partial S}{\partial t}$; the Hamilton-Jacobi equation is satisfied because H is constant. The definition implies that partial derivatives of S and W with respect to $\vec{q}$ are identical, so the Hamilton-Jacobi equation can be rewritten in the form (known as the **restricted Hamilton-Jacobi equation**)

$$H\left(\vec{q}, \frac{\partial W(\vec{q}, \vec{P})}{\partial \vec{q}}\right) = E \tag{2.67}$$

The function W is known as **Hamilton's Characteristic Function**. W can be rewritten in a more physical manner:

$$\begin{aligned} W &= S + E\,t = \int dt\,(L + H) \\ &= \int dt \sum_k p_k\,\dot{q}_k = \int d\vec{q}\cdot\vec{p} \end{aligned} \tag{2.68}$$

which is known as the **abbreviated action**.

W is more valuable than as just another interesting theoretical quantity. The restricted Hamilton-Jacobi equations looks like a canonical transformation of the Hamiltonian by a F_2 generating function because we have an equation where something (E) equals the Hamiltonian. The reason it must be a F_2 function is because the momenta are replaced by $\frac{\partial W}{\partial \vec{q}}$ in the original Hamiltonian.

It should first be realized that the canonical momenta $\vec{P}$ generated by W may not be the same as the $\vec{\alpha}$ generated by S; after all, W is a different function from S. But, clearly, W is very close to S, differing only by the term $E\,t$. One possible choice of the new momenta $\vec{P}$ is to say $P_1 = E$ and leave the remainder unchanged. That is, suppose we had solved for S and found the M constant momenta $\vec{\alpha}$. We have in the relation between S and W another constant E. Since there are only M constants to be specified to define S (neglecting the offset term), E must be some combination of those M constants. That is, of the $M + 1$ constants E and $\vec{\alpha}$, only M are independent. The solution S chooses the $\vec{\alpha}$ as the independent ones and E as the derived one. But we are free to instead make E an independent one and rewrite α_1 in terms of E and the remainder of $\vec{\alpha}$. This is not the *only* choice we could have made, but obviously it is a simple one.

Let us explore whether W does indeed qualify as a generating function and what transformation it generates. Does the above choice of the relation between the momenta $\vec{\alpha}$ from the Hamilton-Jacobi equation and the moment $\vec{P}$ of the restricted Hamilton-Jacobi equation work – does it generate a canonical transformation that makes $\tilde{H}$ simply equal to the canonical momentum E? We can see that the only difference between the transformations generated by S and by W is in P_1 and Q_1. The remaining P_j are left unchanged by definition. The corresponding Q_j are seen to be the same by calculating what they would be if W is indeed a generating function:

$$Q_j = \frac{\partial W}{\partial P_j} = \frac{\partial}{\partial P_j}(S + E\,t) = \frac{\partial S}{\partial P_j} + t\,\frac{\partial E}{\partial P_j} = \beta_j \qquad j \neq 1$$

where $\beta_j = \frac{\partial S}{\partial P_j}$ is the Q_j from our original S transformation and the $\frac{\partial E}{\partial P_j}$ term vanishes for $j \neq 1$ because E is P_1. Since the Q_j and P_j are expressly unchanged for $j > 1$, and neither appears in $\tilde{H} = E$ (the Hamiltonian after the canonical transformation to $(\vec{Q}, \vec{P})$ generated by W), Hamilton's equations are trivially satisfied and the transformation generated by W is canonical in the $j > 1$ coordinates.

What about $j = 1$? The generating function would give us for Q_1:

$$Q_1 = \frac{\partial W}{\partial P_1} = \frac{\partial S}{\partial P_1} + t = \frac{\partial S}{\partial \alpha_1}\frac{\partial \alpha_1}{\partial P_1} + t = \beta_1 \frac{\partial \alpha_1}{\partial P_1} + t \equiv \beta_1' + t$$

The statement $\frac{\partial S}{\partial P_1} = \frac{\partial S}{\partial \alpha_1}\frac{\partial \alpha_1}{\partial P_1}$ is not obvious. The reason it holds is because we know that P_1 can written in terms of α_1 and the other momenta α_j, and vice versa, and that, when all other momenta are held constant, variations in α_1 and P_1 are therefore directly related. Essentially, the derivatives all become one-dimensional and standard chain rule arguments then apply. Does this form for Q_1 show that the transformation is canonical? This form gives $\dot{Q}_1 = 1$. Hamilton's equations in the new coordinates would give $\dot{Q}_1 = \frac{\partial \tilde{H}}{\partial P_1} = 1$, so yes on the first of Hamilton's equations. The other Hamilton's equation, $\dot{P}_1 = -\frac{\partial \tilde{H}}{\partial Q_1}$, is trivially satisfied because both sides vanish: the left side because H is conserved, and the right side because Q_1 does not appear in $\tilde{H}$. So, then, we have shown that the transformation generated by W is indeed canonical.

To summarize: we have found that, when H is conserved and has value E, if $W(q, E, \alpha_2, \ldots, \alpha_M)$ satisfies the equation

$$H\left(\vec{q}, \frac{\partial W(\vec{q}, E, \alpha_2, \ldots, \alpha_M)}{\partial \vec{q}}, t\right) = E$$

then W generates a canonical transformation that transforms H into $\tilde{H} = E$, with the transformation of the coordinates having the generic functional form

$$p_k = \frac{\partial W(\vec{q}, E, \alpha_2, \ldots, \alpha_M)}{\partial q_k} \tag{2.69}$$

$$Q_k = \frac{\partial W(\vec{q}, E, \alpha_2, \ldots, \alpha_M)}{\partial \alpha_k} = \beta_k \qquad k > 1 \tag{2.70}$$

$$Q_1 = \frac{\partial W(\vec{q}, E, \alpha_2, \ldots, \alpha_M)}{\partial E} + t = \beta_1' + t \tag{2.71}$$

that incorporates initial conditions via the equations

$$p_k(t=0) = \left.\frac{\partial W(\vec{q}, E, \alpha_2, \ldots, \alpha_M)}{\partial q_k}\right|_{q(t=0), E, \alpha_2, \ldots, \alpha_M} \tag{2.72}$$

$$\beta_k = \left.\frac{\partial W(\vec{q}, E, \alpha_2, \ldots, \alpha_M)}{\partial \alpha_k}\right|_{q(t=0), E, \alpha_2, \ldots, \alpha_M} \tag{2.73}$$

$$\beta_1' = \left.\frac{\partial W(\vec{q}, E, \alpha_2, \ldots, \alpha_M)}{\partial E}\right|_{q(t=0), E, \alpha_2, \ldots, \alpha_M} \tag{2.74}$$

The distinction between $k = 1$ and $k > 1$ is due to the choice of P_1 as the energy. Note that, in all the equations, the time dependence is nonexistent or very simple. For the more general transformation and generating function S, while the Q_k were all constant, S could have had explicit time dependence, so there might have been an explicit time dependence in the relation between $\vec{q}$, $\vec{\alpha}$, and $\vec{\beta}$. Now, we are guaranteed that there can be such explicit time dependence in only the Q_1 equation. And, since the Q_1 equation is so simple, we can use it to eliminate t completely as an independent variable. We can pick a coordinate (call it q_1 for specificity) as the independent variable and parameterize the evolution of the remaining q_k and all the p_k in terms of it. The resulting relations are called **orbit equations** because they describe the shapes of the particle paths in phase space.

Note that we could have made a more complicated choice of how to relate E to the canonical momenta. Instead of picking a transformation that makes $\tilde{H} = P_1$ a simple function of only the first canonical momentum, we could have chosen $\tilde{H} = \tilde{H}(\vec{P})$ a more complicated (but still time-independent) function of all the momenta. This is a choice, based on what is convenient for the problem. If we had chosen this route, some or all of the equations for Q_k would have had a linear time dependence in them. The system remains fairly simple. This kind of situation is presented in the example on the anisotropic simple harmonic oscillator below.

There is a very nice table in Section 10.3 of Goldstein comparing and contrasting the S and W transformations.

Separability

The abstract discussion above has made no attempt to demonstrate in what cases the equations are actually explicity solvable. One sufficient, but not necessary, condition for solubility is that the Hamilton-Jacobi equation be **separable**; that is, if Hamilton's Principal Function can be written as a sum of M terms, each depending on only one of the original coordinates and time,

$$S(\vec{q}, \vec{\alpha}, t) = \sum_k S_k(q_k, \vec{\alpha}, t) \tag{2.75}$$

then we can explicitly show how to convert the partial differential equation to M ordinary differential equations using **separation of variables**. If S can be written in the above fashion, we are guaranteed that H can also be written this way because H is, theoretically at least, derived from S through L. The Hamilton-Jacobi equation then becomes

$$\sum_k \left[H_k\left(q_k, \frac{\partial S_k}{\partial q_k}, \vec{\alpha}, t\right) + \frac{\partial S_k(q_k, \vec{\alpha}, t)}{\partial t} \right] = 0$$

Since each term in the sum depends on a different coordinate (and the same constants $\vec{\alpha}$), each term must separately vanish, yielding a set of M uncoupled equations

$$H_k\left(q_k, \frac{\partial S_k}{\partial q_k}, \vec{\alpha}, t\right) + \frac{\partial S_k(q_k, \vec{\alpha}, t)}{\partial t} = 0 \tag{2.76}$$

If the H_k are individually conserved (*e.g.*, none have any explicit time dependence), then we can further deduce

$$S_k(q_k, \vec{\alpha}, t) = W_k(q_k, \vec{\alpha}) - \alpha_k\, t$$

and therefore

$$H_k\left(q_k, \frac{\partial W}{\partial q_k}, \vec{\alpha}\right) = \alpha_k \tag{2.77}$$

where we have obviously chosen to take the M constants α_k to be the values of the individually conserved H_k rather than taking $\alpha_1 = \sum_k H_k$ and leaving the remaining α_k to be some linear combination of the H_k. By this choice, the total energy is $E = \sum_k \alpha_k$.

Separability of H in the above fashion is guaranteed if the Hamiltonian meets the **Staeckel Conditions**:

1. The Hamiltonian is conserved (as we already discussed above).
2. The Lagrangian is no more than a quadratic function of the generalized velocities, so the Hamiltonian has the form

$$H = \frac{1}{2}\left(\vec{p} - \vec{a}\right)^T \underline{T}^{-1} \left(\vec{p} - \vec{a}\right) + V(\vec{q})$$

   where the a_k are functions only of the conjugate coordinate, $a_k = a_k(q_k)$ only, $\underline{T}$ is a symmetric, invertible matrix, and $V(\vec{q})$ is a potential energy function that depends only on the coordinates.
3. The potential energy can be written in the form

$$V(\vec{q}) = \sum_k \frac{V(q_k)}{\underline{T}_{kk}}$$

4. And the final, inscrutable condition: If we define a matrix $\underline{\phi}$ by

$$\sum_l \delta_{kl}\, \underline{\phi}^{-1}_{kl} = \frac{1}{\underline{T}_{kk}}$$

   with

$$\frac{\partial W_k}{\partial q_k} - a_k = \sum_{lm} 2\, \delta_{kl}\, \phi_{lm}\, \gamma_m$$

   where $\vec{\gamma}$ is an unspecified constant vector. The diagonal elements of $\underline{\phi}$ and $\underline{\phi}^{-1}$ may depend only on the associated coordinate.

We will not attempt to prove these conditions; they are proven in Appendix D of the second edition of Goldstein.

Examples

Example 2.16: Simple harmonic oscillator

The simple harmonic oscillator Hamiltonian is (using the same form as used in Example 2.13):

$$H = \frac{1}{2}\left(p^2 + \omega^2\, q^2\right) \equiv E$$

where we have explicitly written the conserved value of H as E. The Hamilton-Jacobi equation for this Hamiltonian is

$$\frac{1}{2}\left[\left(\frac{\partial S}{\partial q}\right)^2 + \omega^2\, q^2\right] + \frac{\partial S}{\partial t} = 0$$

where we have made the substitution $p = \frac{\partial S}{\partial q}$ in keeping with the assumption that S is a F_2 generating function. Since H is conserved, we may proceed to the Hamilton's Characteristic Function, writing $\frac{\partial S}{\partial t} = \alpha$ a constant:

$$\frac{1}{2}\left[\left(\frac{\partial W}{\partial q}\right)^2 + \omega^2\,q^2\right] = \alpha$$

Obviously, $\alpha = E$ since the left side of the equation is the Hamiltonian. The above equation is directly integrable since it is now a differential equation in q only:

$$\begin{aligned}\frac{\partial W}{\partial q} &= \sqrt{2\,E - \omega^2\,q^2}\\ W &= \sqrt{2\,E}\int dq\,\sqrt{1 - \frac{\omega^2\,q^2}{2\,E}}\\ S &= -E\,t + \sqrt{2\,E}\int dq\,\sqrt{1 - \frac{\omega^2\,q^2}{2\,E}}\end{aligned}$$

We neglect to perform the integration (which is entirely doable via trigonometric substitution, resulting in a cos function) because we only need the partial derivatives of S. We have already evaluated the one constant canonical momentum, so let us obtain the corresponding constant β for the linearly evolving coordinate using Equation 2.71:

$$\begin{aligned}t + \beta &= \frac{\partial W}{\partial \alpha} = \frac{\partial W}{\partial E}\\ &= \sqrt{\frac{1}{2\,E}}\int \frac{dq}{\sqrt{1 - \frac{\omega^2\,q^2}{2\,E}}}\\ &= \frac{1}{\omega}\,\arcsin\frac{\omega\,q}{\sqrt{2\,E}}\end{aligned}$$

which is easily inverted to give

$$q = \frac{\sqrt{2\,E}}{\omega}\,\sin\left(\omega\,t + \phi\right)$$

where $\phi = \omega\,\beta$ is just a rewriting of the constant. We can now obtain p from Equation 2.70:

$$\begin{aligned}p &= \frac{\partial W}{\partial q} = \sqrt{2\,E}\sqrt{1 - \frac{\omega^2\,q^2}{2\,E}}\\ &= \sqrt{2\,E}\,\cos\left(\omega\,t + \phi\right)\end{aligned}$$

Finally, we need to connect the constants E and ϕ (α and β in the formal derivation) with the initial conditions. We could go back to Equations 2.65 and 2.65, but it is easier to simply make use of the solutions directly. From the Hamiltonian, we see

$$E = \frac{1}{2}\left(p_0^2 + \omega^2\,q_0^2\right)$$

and from the solutions for $q(t)$ and $p(t)$ we see

$$\tan\phi = \omega\,\frac{q_0}{p_0}$$

Hamilton's Principal Function in fact generates the same canonical transformation as we saw in Example 2.13, converting from position and momentum to energy and phase. The energy is conserved and the phase evolves linearly with time. If we want, we can recover Hamilton's Principal Function explicitly by substituting the solution into our integral form for S and integrating:

$$\begin{aligned}
S &= -E\,t + \sqrt{2\,E}\int dq\,\sqrt{1-\frac{\omega^2\,q^2}{2\,E}} \\
&= -E\,t + \sqrt{2\,E}\int \left[d\left(\frac{\sqrt{2\,E}}{\omega}\sin\left(\omega\,t+\phi\right)\right)\right]\sqrt{1-\sin^2\left(\omega\,t+\phi\right)} \\
&= -E\,t + 2\,E\int dt\,\cos^2\left(\omega\,t+\phi\right) \\
&= 2\,E\int dt\left[\cos^2\left(\omega\,t+\phi\right)-\frac{1}{2}\right]
\end{aligned}$$

where notice that we not only had to substitute for q but for dq also. If we calculate the Lagrangian directly, we have

$$\begin{aligned}
L &= \frac{1}{2}\left(p^2-\omega^2\,q^2\right) \\
&= E\left(\cos^2\left(\omega\,t+\phi\right)-\sin^2\left(\omega\,t+\phi\right)\right) \\
&= 2\,E\left[\cos^2\left(\omega\,t+\phi\right)-\frac{1}{2}\right]
\end{aligned}$$

and we see that $S = \int dt\,L$ explicitly.

Example 2.17: Two-dimensional anisotropic harmonic oscillator

An anisotropic two-dimensional harmonic oscillator has different spring constants (and therefore different characteristic frequencies) in the two dimensions. It thus does not trivially separate into cylindrical coordinates. The Hamiltonian is

$$H = \frac{1}{2}\left(p_x^2+p_y^2+\omega_x^2\,x^2+\omega_y^2\,y^2\right) \equiv E$$

The Hamiltonian is clearly separable in Cartesian coordinates, so we are led to a Hamilton's Principal Function of the form

$$S(x,y,\alpha_x,\alpha_y,t) = W_x(x,\alpha_x)+W_y(y,\alpha_y)-(\alpha_x+\alpha_y)\,t$$

where $E = \alpha_x + \alpha_y$. Here we have chosen to do the problem symmetrically in the two momenta rather than pick one to be the energy. Since the system is separable, we may go directly to Equation 2.77:

$$\begin{aligned}
\frac{1}{2}\left[\left(\frac{\partial W_x}{\partial x}\right)^2+\omega_x^2\,x^2\right] &= \alpha_x \\
\frac{1}{2}\left[\left(\frac{\partial W_y}{\partial y}\right)^2+\omega_y^2\,y^2\right] &= \alpha_y
\end{aligned}$$

Each equation is just a 1-dimensional oscillator equation, so the solution is

$$\begin{aligned}
x &= \frac{\sqrt{2\,\alpha_x}}{\omega_x}\,\sin\left(\omega_x\,t + \phi_x\right)\\
p_x &= \sqrt{2\,\alpha_x}\,\cos\left(\omega_x\,t + \phi_x\right)\\
y &= \frac{\sqrt{2\,\alpha_y}}{\omega_y}\,\sin\left(\omega_y\,t + \phi_y\right)\\
p_y &= \sqrt{2\,\alpha_y}\,\cos\left(\omega_y\,t + \phi_y\right)
\end{aligned}$$

Example 2.18: Isotropic two-dimensional harmonic oscillator

Here, $\omega_x = \omega_y$ so we write the Hamiltonian in polar coordinates (r, θ):

$$H = \frac{1}{2}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \omega^2\,r^2\right) \equiv E$$

where $p_\theta = r^2\,\dot{\theta}$ is the canonical momentum conjugate to θ. The problem is not trivially separable because the second term depends on both r and p_θ, mixing the two coordinates. However, because the Hamiltonian is cyclic in θ, we are assured that $p_\theta = \alpha_\theta$ is constant. This is the condition we need to be able to separate variables: Hamilton's Principal Function must be writeable as a sum of terms $S_k(q_k, \vec{\alpha}, t)$, which we can indeed do:

$$S(r, \theta, E, \alpha_\theta) = W_r(r, E, \alpha_\theta) + W_\theta(\theta, E, \alpha_\theta) - E\,t$$

where here we do not try to have symmetric constants because there is no symmetry between r and θ. The reduced Hamilton-Jacobi equation is

$$\frac{1}{2}\left[\left(\frac{\partial W_r}{\partial r}\right)^2 + \frac{\alpha_\theta^2}{r^2} + \omega^2\,r^2\right] = E$$

There is no equation for θ because there are no terms in H that depend on θ. We can trivially obtain W_θ by making use of

$$\begin{aligned}
\frac{\partial W_\theta}{\partial \theta} &= p_\theta = \alpha_\theta\\
W_\theta &= \alpha_\theta\,\theta
\end{aligned}$$

At this point, we are essentially done with the formal part, and we need only to solve the above differential equation for r mathematically. But that equation is not trivially integrable, so we pursue a different path (with some foreknowledge of the solution!). The solution is no doubt going to involve simple harmonic motion in the two dimensions, so we try

$$\begin{aligned}
x &= \frac{\sqrt{2\,\alpha}}{\omega}\,\sin\left(\omega\,t + \phi\right)\\
p_x &= \sqrt{2\,\alpha}\,\cos\left(\omega\,t + \phi\right)\\
y &= \frac{\sqrt{2\,\alpha}}{\omega}\,\sin\omega\,t\\
p_y &= \sqrt{2\,\alpha}\,\cos\omega\,t
\end{aligned}$$

where we include a phase factor for x only because there can be only one undetermined constant of integration for this first-order differential equation. The resulting polar and angular coordinate and momenta solutions are

$$\begin{aligned} r &= \frac{\sqrt{2\,E}}{\omega}\sqrt{\sin^2(\omega\,t+\phi)+\sin^2\omega\,t} \\ p_r &= \dot{r} \\ \theta &= \arctan\left[\frac{\sin\omega\,t}{\sin(\omega\,t+\phi)}\right] \\ p_\theta &= r^2\,\dot{\theta} = \alpha_\theta \end{aligned}$$

For completeness, let us calculate p_θ explicitly from θ to see what form it has.

$$\begin{aligned} p_\theta &= r^2\,\dot{\theta} \\ &= \frac{2\,E}{\omega^2}\left[\sin^2(\omega\,t+\phi)+\sin^2\omega\,t\right]\frac{d}{dt}\arctan\left[\frac{\sin\omega\,t}{\sin(\omega\,t+\phi)}\right] \\ &= \frac{2\,E}{\omega^2}\left[\sin^2(\omega\,t+\phi)+\sin^2\omega\,t\right]\left\{1+\frac{\sin^2\omega\,t}{\sin^2(\omega\,t+\phi)}\right\}^{-1}\frac{d}{dt}\left[\frac{\sin\omega\,t}{\sin(\omega\,t+\phi)}\right] \\ &= \frac{2\,E}{\omega^2}\,\frac{\sin^2(\omega\,t+\phi)+\sin^2\omega\,t}{1+\frac{\sin^2\omega\,t}{\sin^2(\omega\,t+\phi)}}\left\{\frac{\omega\,\cos\omega\,t}{\sin(\omega\,t+\phi)}-\frac{\sin\omega\,t}{\sin^2(\omega\,t+\phi)}\,\omega\,\cos(\omega\,t+\phi)\right\} \\ &= \frac{2\,E}{\omega}\left[\cos\omega\,t\,\sin(\omega\,t+\phi)-\sin\omega\,t\cos(\omega\,t+\phi)\right] \\ &= \frac{2\,E}{\omega}\,\sin\phi \end{aligned}$$

which is very nice. If we consider $\phi = 0$, so the two oscillators are perfectly in phase, then the motion is along a straight line through the origin:

$$\begin{aligned} r &= \frac{2\sqrt{E}}{\omega}\,|\sin\omega\,t| \\ p_r &= \sqrt{2\,E}\cos\omega\,t \\ \theta &= \arctan 1 = \frac{\pi}{4} \\ p_\theta &= 0 \end{aligned}$$

If we consider $\phi = \frac{\pi}{2}$, the two oscillators are $\pi/2$ out of phase and the motion is circular:

$$\begin{aligned} r &= \frac{\sqrt{2\,E}}{\omega}\sqrt{\cos^2\omega\,t+\sin^2\omega\,t} = \frac{\sqrt{2\,E}}{\omega} \\ p_r &= 0 \\ \theta &= \omega\,t \\ p_\theta &= \frac{2E}{\omega} = r^2\,\omega \end{aligned}$$

as one would expect for perfect circular motion.

Additional Examples: See Hand and Finch Section 6.5 – there are a couple of gravitational examples.

Deriving Action-Angle Variables via the Hamilton-Jacobi Formalism

Earlier, we introduced the concept of action-angle variables for 1-dimensional systems with a conserved Hamiltonian by using F_1-type and F_2-type generating functions to transform to coordinates (ψ, I) such that I is constant and ψ evolves linearly with time. This is obviously quite analogous to the Hamilton-Jacobi formalism, so it is interesting to repeat the derivation using those results.

We have a 1-dimensional system for which the Hamiltonian is conserved. We can then certainly use the Hamilton-Jacobi formalism, specializing to the conserved Hamiltonian case. That formalism gives us the new generating function, Hamilton's Characteristic Function (see Equation 2.68)

$$W = \int p\, dq$$

and tells us that the transformation generated by W makes the Hamiltonian identically equal to the constant momentum variable α_1 and yields a linearly evolving coordinate $Q_1 = t + \beta_1$ where β_1 is set by initial conditions. Now, since I is not identically equal to E, this exact equation for Q_1 does not hold for ψ. But we can find ψ. First, the explicit form of ψ is given by the generating function, which yields the relation

$$\psi = \frac{\partial W}{\partial I} = \frac{\partial}{\partial I} \int p\, dq$$

Where does I come in? Remember that we wrote earlier $p = p(q, \alpha_1)$. We now have two constants I and α_1 when we know there is only one independent constant, so we can choose I as that constant and rewrite α_1 in terms of I. So

$$\psi = \frac{\partial}{\partial I} \int p(q, I)\, dq$$

The time evolution of ψ is found by Hamilton's equations

$$\dot{\psi} = \frac{\partial H(I)}{\partial I} \equiv \omega$$

which is just some constant determined by the particular functional form of $H(I)$.

Hamilton-Jacobi Theory, Wave Mechanics, and Quantum Mechanics

Think about S as a surface in configuration space (the M-dimensional space of the system coordinates $\vec{q}$). The Hamilton-Jacobi equation finds a function S such that the momenta are $p_k = \frac{\partial S}{\partial q_k}$, which can be written $\vec{p} = \vec{\nabla}_q S$. The gradient of a function gives a vector field normal to surfaces over which the function value is constant. So the momentum vector is normal to surfaces of constant S. The momentum vector is also tangent to the particle trajectory in configuration space. So the particle trajectories are always normal to the surfaces of constant S. For conserved Hamiltonians, where we write $S = W - E\,t$, the shapes of the surfaces are given by Hamilton's Characteristic Function W and they move linearly in time due to the $-E\,t$ term.

In optics, a classical electromagnetic propagates in a similar fashion. There are surfaces of constant phase ϕ. These are like the surfaces of constant S. The wave propagates by following the normal to the surfaces of constant phase just as the mechanical system follows the trajectories perpendicular to the surfaces of constant S. Geometrical optics rays are analogous to the particle trajectories, as these rays are always perpendicular to surfaces of constant phase. The geometrical optics wavevector $\vec{k}$ is the obvious analogy to the momentum $\vec{p}$.

Thus, in the way that a classical wave's phase advances linearly in time – the surfaces of constant phase propagate linearly outward according to the wave's phase velocity – similarly the surfaces of constant S propagate forward at a speed determined by the total energy E. The action integral S appears to have the characteristics of a wave phase function ϕ.

In the pre-quantum era, this analogy was little more than a mathematical curiosity, perhaps providing some practical benefit by allowing the flow of solutions between mechanics and optics. The power of the analogy becomes clear only when we consider how it relates to quantum mechanics.

In quantum mechanics, a particle is described not by a simple trajectory $\vec{q}(t)$, but rather by a wave function $\psi(\vec{x}, t)$ that can have nonzero value at all points in space and time. If we let $\rho(\vec{x}, t) = |\psi(\vec{x}, t)|^2$, then ρ describes the probability of finding the particle at position $\vec{x}$ at time t. The classical trajectory is given by $\vec{q}(t) = \int d^3x\, \vec{x}\, |\psi(\vec{x}, t)|^2$; *i.e.*, it is the first moment of the wave function. In one dimension, the wave function satisfies Schrodinger's equation,

$$-\frac{\hbar^2}{2\,m}\frac{\partial^2 \psi}{\partial x^2} + V(x)\,\psi = i\,\hbar\,\frac{\partial \psi}{\partial t}$$

where $V(x)$ is the classical potential in which the particle moves. The wave function clearly must have the form $\psi(x, t) = \sqrt{\rho(x, t)}\, e^{i\,\phi(x,t)}$. Given the analogy between S and a classical wave's phase ϕ above, let us in fact assume $\phi = S/\hbar$. As $\hbar \to 0$, the phase of the wave oscillates faster and faster, which we might expect would hide the wave characteristics (waves are most "wave-like" in the long-wavelength limit) as is necessary in this limit. If we rewrite $\psi(x, t) = \sqrt{\rho(x, t)}\, e^{\frac{i}{\hbar}\,S(x,t)}$, Schrodinger's equation becomes

$$-\frac{\hbar^2}{2\,m}\left\{\frac{\partial^2\sqrt{\rho}}{\partial x^2} + \frac{2\,i}{\hbar}\frac{\partial\sqrt{\rho}}{\partial x}\frac{\partial S}{\partial x} - \frac{1}{\hbar^2}\sqrt{\rho}\left(\frac{\partial S}{\partial x}\right)^2 + \frac{i}{\hbar}\sqrt{\rho}\,\frac{\partial^2 S}{\partial x^2}\right\} e^{\frac{i}{\hbar}\,S} + V(x)\,\sqrt{\rho}\,e^{\frac{i}{\hbar}\,S}$$
$$= i\,\hbar\left\{\frac{\partial\sqrt{\rho}}{\partial t} + \frac{i}{\hbar}\sqrt{\rho}\,\frac{\partial S}{\partial t}\right\} e^{\frac{i}{\hbar}\,S}$$

The equation contains terms up to second order in $\hbar$; neglect all terms containing $\hbar$ or $\hbar^2$. This leaves (canceling the common $\sqrt{\rho}\,e^{\frac{i}{\hbar}\,S}$)

$$\frac{1}{2\,m}\left(\frac{\partial S}{\partial x}\right)^2 + V(x) + \frac{\partial S}{\partial t} = 0$$

We thus recover the Hamilton-Jacobi equation for a particle propagating in a potential $V(x)$.

3.1 The Simple Harmonic Oscillator

The linear simple harmonic oscillator (SHO) is the foundation of the theory of oscillations. We discuss equilibria in physical systems and how small oscillations about equilibria can in most cases be described by the SHO equation. We investigate the effect of damping and driving forces on the SHO, using the opportunity to introduce Green's functions.

We follow Hand and Finch Chapter 3 for the most part, though we make small changes here and there. Most intermediate mechanics texts will have a similar discussion of the basics, though most will not discuss Green's functions.

3.1.1 Equilibria and Oscillations

Types of Equilibria

Recall in Section 1.1.3 that the equilibrium points of a conservative system (one with a potential energy function) can be found by requiring $\vec{\nabla} U = 0$, and that the stability of the equilibria is determined by the sign of the second derivatives. In one dimension, the types of equilibria are

- **stable equilibrium**: $d^2U/dx^2 > 0$
- **unstable equilibrium**: $d^2U/dx^2 < 0$
- **saddle point**: $d^2U/dx^2 = 0$

With rheonomic systems, the stability of an equilibrium point may change in time.

Equilibria from the Lagrangian Point of View

Consider a Taylor expansion around an equilibrium point $(q_0, \dot{q}_0)$ of an arbitrary 1-dimension Lagrangian to second order in the generalized coordinates and velocities:

$$L \approx A + B\,q + C\,\dot{q} + D\,q^2 + E\,q\,\dot{q} + F\,\dot{q}^2$$

We neglect A since it is a constant offset and does not affect the dynamics. The constants are given by assorted partial derivatives:

$$B = \left.\frac{\partial L}{\partial q}\right|_{(q_0,\dot{q}_0)} \qquad C = \left.\frac{\partial L}{\partial \dot{q}}\right|_{(q_0,\dot{q}_0)}$$

$$D = \frac{1}{2}\left.\frac{\partial^2 L}{\partial q^2}\right|_{(q_0,\dot{q}_0)} \qquad E = \left.\frac{\partial^2 L}{\partial \dot{q}\,\partial q}\right|_{(q_0,\dot{q}_0)} \qquad F = \frac{1}{2}\left.\frac{\partial^2 L}{\partial \dot{q}^2}\right|_{(q_0,\dot{q}_0)}$$

$B = 0$ in order for $(q_0, \dot{q}_0)$ to be an equilibrium position. The Euler-Lagrange equation for the system is

$$\frac{d}{dt}\left(C + E\,q + 2\,F\,\dot{q}\right) = 2\,D\,q + E\,\dot{q}$$

$$\ddot{q} - \frac{D}{F}\,q = 0$$

We would have found the same equation of motion had we started with

$$L \approx D\,q^2 + F\,\dot{q}^2$$

This is just the simple harmonic oscillator Lagrangian with characteristic frequency

$$\omega^2 \equiv -\frac{D}{F}$$

Now, let's rescale the time units to get rid of the coefficients: define $\beta = \sqrt{\left|\frac{F}{D}\right|}$ and define a new time coordinate τ by $t = \beta\,\tau$. The equation of motion and Lagrangian becomes

$$\ddot{q} \pm q = 0 \qquad L = \frac{1}{2}\left(\dot{q}^2 \mp q^2\right) \tag{3.1}$$

where the sign is the opposite of the sign of $\frac{D}{F}$. For reasonable Lagrangians, F is the coefficient of the kinetic energy term and thus $F > 0$ holds.

Restricting to conservative, scleronomic systems, we are assured that the sign of D is set by the sign of the second derivative of the potential energy, so stability is, as we found before, determined by the shape of the potential energy function:

$$\begin{array}{llll} \text{stable:} & \ddot{q} + q = 0 & D < 0 & \frac{\partial^2 U}{\partial q^2} > 0 \\ \text{unstable:} & \ddot{q} - q = 0 & D > 0 & \frac{\partial^2 U}{\partial q^2} < 0 \end{array}$$

Thus, we see that working from the Lagrangian perspective yields, as we would expect, the same conditions for stable and unstable equilibria as we found from the Newtonian perspective. More importantly, we see the central reason for studying the SHO in detail: *the Taylor expansion of (almost) any scleronomic, conservative, one-dimensional system about a stable equilibrium looks like a simple harmonic oscillator.* We will not show it here, but this conclusion extends to multi-dimensional systems, too.

The Hamiltonian

Finally, we note that the above Taylor expansion also lets us write an approximate Hamiltonian for the system. Starting from the original L, we have

$$\begin{aligned} p &\approx \frac{\partial L}{\partial \dot{q}} = C + E\,q + 2\,F\,\dot{q} \\ H &= p\,\dot{q} - L \\ &\approx C\,\dot{q} + E\,q\,\dot{q} + 2\,F\,\dot{q}^2 - A - B\,q - C\,\dot{q} - D\,q^2 - E\,q\,\dot{q} - F\,\dot{q}^2 \\ &= F\,\dot{q}^2 - D\,q^2 \end{aligned}$$

or, after rescaling

$$H = \frac{1}{2}\left(\dot{q}^2 \pm q^2\right) \tag{3.2}$$

where the sign again determines the stability, + is stable, − is unstable. Since we earlier demonstrated that the C and E terms do not affect the dynamics, we are free to drop them and redefine the momentum simply as

$$p = 2\,F\,\dot{q}$$

3.1.2 Solving the Simple Harmonic Oscillator

Characteristics of the Equation of Motion

The equation of motion for the stable simple harmonic oscillator has been written

$$\ddot{q} + q = 0$$

This is a linear, second-order differential equation with constant coefficients. It is:

- **homogeneous**: (the right side vanishes), so solutions can be scaled by a constant and still satisfy the equation.
- **linear**: Because the equation is linear in q and its time derivatives, a linear combination $a\, q_1(t) + b\, q_2(t)$ of two solutions is also a solution.

Conventional Solution

You are no doubt well aware that the general solution of this equation of motion is a sinusoid, which can be written in two forms:

$$a(t) = A\, \sin(t + \phi) = A' \cos t + B' \sin t$$

where the two sets of coefficients are related to each other and the initial conditions by

$$A\, \sin\phi = A' = q(0) \qquad A\, \cos\phi = B' = \dot{q}(0)$$

The period of oscillation is $T = 2\,\pi$, the frequency is $\nu = 1/2\,\pi$, and the angular frequency is $\omega = 1$. For arbitrary $\omega \neq 1$, these become $\nu = \omega/2\,\pi$ and $T = 2\,\pi/\omega$.

Complex Solution

Given the sinusoidal nature of the solution, we could equally well have written it as the real part of a complex function:

$$q_c(t) = \mathcal{A}_c\, e^{it} = A_c\, e^{i\phi}\, e^{it} = A_c\, [\cos(t + \phi) + i\, \sin(t + \phi)] \tag{3.3}$$

where $\mathcal{A}_c = A_c\, e^{i\phi}$ (and thus $A_c = |\mathcal{A}_c|$). The initial conditions are

$$\begin{aligned} q(0) &= \mathcal{R}[q_c(0)] = A_c\, \cos\phi = \mathcal{R}[\mathcal{A}_c] \\ \dot{q}(0) &= \mathcal{R}[\dot{q}_c(0)] = -A_c\, \sin\phi = -\mathcal{I}[\mathcal{A}_c] \end{aligned}$$

or, equivalently,

$$\mathcal{A}_c = q(0) - i\, \dot{q}(0)$$

where $\mathcal{R}[\,]$ and $\mathcal{I}[\,]$ take the real and imaginary part of a complex number. When a physical solution is needed, we simply take $\mathcal{R}[q(t)]$.

We note that a shift of the time origin by t_0 is now just a simple phase factor,

$$\mathcal{A}_c = [q(t_0) - i\, \dot{q}(t_0)]\, e^{-it_0}$$

Note the sign in the phase factor.

3.1.3 The Damped Simple Harmonic Oscillator

Most simple harmonic oscillators in the real world are damped – mechanical oscillators, electrical oscillators, etc. Thus, the damped simple harmonic oscillator is the next important system to study.

Introducing Damping

We assume that a damping force linear in velocity is applied to the harmonic oscillator. For a mechanical oscillator, this could be a frictional force. For an electrical oscillator, this could be a resistive element. Regardless, the frictional force is assumed to be of the form

$$F_{damp} = -\frac{\dot{q}}{Q}$$

The dimensionless constant Q is called the **quality factor** and is ubiquitous in the description of oscillatory systems with damping.

It is nontrivial to introduce frictional forces via the Lagrangian formalism. Since we have already written an equation of motion, it is straightforward to include the damping term via Newtonian mechanics. The equation of motion of the **damped simple harmonic oscillator** is

$$\ddot{q} + \frac{\dot{q}}{Q} + q = 0 \tag{3.4}$$

This is obviously not the most generic frictional force that one can think of. It is, however, the kind that is encountered in most damped electrical circuits and thus is of great practical interest.

It is instructive to go back to a dimensional version of the equation to understand the physical meaning of Q. The equation of a motion of a mechanical oscillator subject to a linear damping force is

$$\begin{aligned} m\,\ddot{x} + b\,\dot{x} + k\,x &= 0 \\ \ddot{x} + \frac{b}{m}\,\dot{x} + \frac{k}{m}\,x &= 0 \end{aligned}$$

If time is rescaled by $\beta = \sqrt{\frac{m}{k}} = \omega^{-1}$ (*i.e.*, $t = \beta\,\tau$), then $Q^{-1} = \frac{b/m}{\omega} = \frac{b}{\sqrt{k\,m}}$. Q decreases as the damping constant increases, and Q increases as the spring constant or mass increase relative to the damping constant (both of which will tend to make the damping constant less important, either because there is a larger force for a given displacement or because the moving mass has more inertia).

The equation of motion of a series LRC oscillator circuit (where q represents the charge on the capacitor and $\dot{q}$ is the current flowing in the circuit) is[1]

$$\begin{aligned} L\,\ddot{q} + \dot{q}\,R + \frac{q}{C} &= 0 \\ \ddot{q} + \frac{\dot{q}}{L/R} + \frac{1}{LC}\,q &= 0 \\ \ddot{q} + \frac{\dot{q}}{\tau_{damp}} + \omega^2\,q &= 0 \end{aligned}$$

[1] The choice of $\tau_{damp} = L/R$ instead of $\tau_{damp} = RC$ is determined by the correspondence $b \leftrightarrow R$, $m \leftrightarrow L$.

When we rescale time by using $t = \beta\,\tau$ with $\beta = \sqrt{L\,C} = \omega^{-1}$, we obtain

$$\ddot{q} + \frac{\dot{q}}{\omega\,\tau_{damp}} + q = 0$$

giving $Q = \omega\,\tau_{damp} = 2\,\pi\,\tau_{damp}/T$ where $T = 2\,\pi/\omega$ is the oscillation period. That is, the quality factor can be thought of as the ratio of the damping timescale to the oscillation timescale; if the damping timescale is many oscillation timescales, then the resonator has high Q.

A note on units: for reference, we rewrite the various equations with all the ωs in place so the units are clear.

$$F_{restore} = -k\,x$$

$$F_{damp} = -b\,\dot{x} = \frac{m\,\omega}{Q}\,\dot{x} = \frac{\sqrt{k\,m}}{Q}\,\dot{x}$$

$$\ddot{x} + \frac{\omega}{Q}\,\dot{x} + \omega^2\,x = 0$$

Correspondence with Other Textbooks

Unfortunately, different textbooks define Q in different ways. We list here the three possible conventions so the reader may understand the correspondence between the results derived here and those in other textbooks. The three choices essentially consist of whether Q is defined relative to the natural frequency, the damped characteristic frequency, or the amplitude resonant frequency. Explicitly, the three choices are (using the example of a mechanical resonator with frictional damping force with coefficient b):

- Relative to natural frequency:

$$Q = \frac{\omega}{b/m} \qquad \text{where} \qquad \omega = \sqrt{\frac{k}{m}} = \text{natural frequency}$$

  This can also be written

$$Q = \omega\,\tau_{damp}$$

  This is the convention used in these notes.

- Relative the damped oscillation frequency:

$$Q = \frac{\omega'}{b/m} \qquad \text{where} \qquad \omega' = \sqrt{\frac{k}{m} - \frac{1}{4}\left(\frac{b}{m}\right)^2} = \omega\sqrt{1 - \frac{1}{4}\left(\frac{b/m}{\omega}\right)^2}$$

  which can be written

$$Q = \omega'\,\tau_{damp}$$

- Relative to the amplitude resonant frequency (defined later in Section 3.1.5):

$$Q = \frac{\omega_r}{b/m} \qquad \text{where} \qquad \omega_r = \sqrt{\frac{k}{m} - \frac{1}{2}\left(\frac{b}{m}\right)^2} = \omega\sqrt{1 - \frac{1}{2}\left(\frac{b/m}{\omega}\right)^2}$$

  which can be written

$$Q = \omega_r \tau_{damp}$$

  This is the convention used in Thornton.

It seems obvious that the first equation is most sensible because it discusses the ratio of damping and natural timescales; the other definitions refer to the ratio of the damping to the damped characteristic or the resonant frequency timescales, both of which already incorporate a damping correction; they are thus "moving targets." Moreover, the first definition results in Q giving in very simple form the amplitude and energy decay times. And, of course, in the limit of high Q, the three definitions become identical.

Damped Solutions

Returning to the formal problem: we can try to solve via an exponential solution of the kind tried in the undamped case, $q = e^{i\alpha t}$. One ends up with the equation

$$\left(-\alpha^2 + i\,\frac{\alpha}{Q} + 1\right) e^{i\alpha t} = 0$$

which is solved by the algebraic relation

$$\alpha = \frac{i}{2\,Q} \pm \sqrt{1 - \frac{1}{4\,Q^2}}$$

The solution will be

$$q(t) = \exp\left(-\frac{t}{2\,Q}\right)\,\exp\left(\pm i\,t\,\sqrt{1 - \frac{1}{4\,Q^2}}\right)$$

There is always a decay term. Depending on the sign of the discriminant, the second term may be oscillatory or decaying or it may simply give 1. We examine these cases separately:

- **Underdamped**: $Q > \frac{1}{2}$. In this case, the discriminant is positive and we obtain oscillatory motion from the second term. The complex solution is given by

$$q_c(t) = \mathcal{A}_c\,\exp\left(-\frac{t}{\tau_d}\right)\,\exp\left(\pm i\,\omega'\,t\right) \qquad \tau_d \equiv 2\,Q \qquad \omega' \equiv \sqrt{1 - \frac{1}{4\,Q^2}} \tag{3.5}$$

  The oscillation frequency is shifted. The decay time is simply related to the quality factor. The physical version of the solution is found by taking the real part,

$$q(t) = A_c\,\exp\left(-\frac{t}{\tau_d}\right)\,\cos\left(\omega'\,t + \phi\right) \qquad \mathcal{A}_c = |\mathcal{A}_c|\,e^{i\phi}$$

  Note that the shift in the frequency will be negligible for $Q \gg \frac{1}{2}$. The shift is quadratically suppressed; for large Q, Taylor expansion of the radical gives

$$\omega' \approx 1 - \frac{1}{8\,Q^2}$$

- **Overdamped**: $Q < \frac{1}{2}$. In this case, the discriminant becomes negative and the second term also is decaying. There are actually two possible decay times due to the sign freedom for the radical. The general solution is

$$q(t) = A\,\exp\left(-\frac{t}{\tau_{d,+}}\right) + B\,\exp\left(-\frac{t}{\tau_{d,-}}\right) \qquad \tau_{d,\mp}^{-1} = \frac{1}{2\,Q} \pm \sqrt{\frac{1}{4\,Q^2} - 1} \tag{3.6}$$

The inversion of the sign in subscripting the solution is so that the + subscript goes with the larger time constant. We refer to the decay constants as times (unlike Hand and Finch) because it's more intuitive. We have two possible decay times, one "fast" and one "slow". In the extreme limit of $Q \ll \frac{1}{2}$, the two solutions are extremely different:

$$\tau_{d,+} \approx \frac{1}{Q} \qquad \tau_{d,-} \approx Q$$

- **Critically damped**: $Q = \frac{1}{2}$. The discriminant vanishes and the two solutions become degenerate. The degenerate time constant is $\tau_d = 2\,Q = 1$ (*i.e.*, the damping time becomes $T/2\,\pi$ where T is the undamped oscillation period). However, we must have another solution because the differential equation is second-order. We can motivate a second solution by considering the limit $Q \to \frac{1}{2}$. The two decay time constants are

$$\tau_{d,\mp}^{-1} = 1 \pm \sqrt{\epsilon}$$

where $|\epsilon| \ll 1$ is the small but nonzero radical. Consider two superpositions of the two solutions

$$\begin{aligned} q_1(t) &= \exp\left(-\frac{t}{\tau_{d,+}}\right) + \exp\left(-\frac{t}{\tau_{d,-}}\right) \\ &= \exp(-t)\left[\exp\left(t\sqrt{\epsilon}\right) + \exp\left(-t\sqrt{\epsilon}\right)\right] \\ &\approx 2\,\exp(-t) \\ q_2(t) &= \exp\left(-\frac{t}{\tau_{d,+}}\right) - \exp\left(-\frac{t}{\tau_{d,-}}\right) \\ &= \exp(-t)\left[\exp\left(t\sqrt{\epsilon}\right) - \exp\left(-t\sqrt{\epsilon}\right)\right] \\ &\approx \exp(-t)\left[1 + t\sqrt{\epsilon} - 1 + t\sqrt{\epsilon}\right] \\ &= 2\,t\sqrt{\epsilon}\,\exp(-t) \end{aligned}$$

The first superposition is just the solution we already had. The second superposition gives us a valid second solution. One can confirm that it is a valid solution for $Q = \frac{1}{2}$:

$$\left(\frac{d^2}{dt^2} + 2\,\frac{d}{dt} + 1\right) t\,\exp(-t) = (-2 + t + 2\,(1-t) + t)\,\exp(-t) = 0$$

So, we have that in the case $Q = \frac{1}{2}$, the generic solution is

$$q(t) = A\,\exp(-t) + B\,t\,\exp(-t)$$

For illustrations of the solutions, see Hand and Finch Figure 3.3. Note the distinct nature of the two solutions in all cases.

When it is desired to return to unscaled time for the above solutions, all occurrences of t should be replaced by $\omega\,t$ and all occurrences of ω' should be replaced by ω'/ω.

Energy Decay

The physical interpretation of Q is best seen in the underdamped case. There is a clear decay of the oscillation amplitude with time constant $\tau_d = 2\,Q$. Since the energy in the

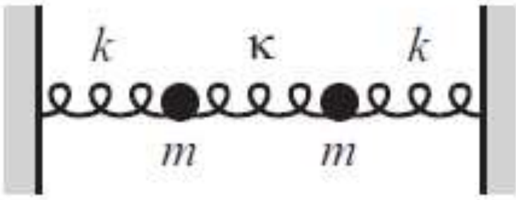
Matthew Schwartz

Lecture 3:
Coupled oscillators

1 Two masses

To get to waves from oscillators, we have to start coupling them together. In the limit of a large number of coupled oscillators, we will find solutions while look like waves. Certain features of waves, such as resonance and normal modes, can be understood with a finite number of oscillators. Thus we start with two oscillators.

Consider two masses attached with springs



(1)

Let's say the masses are identical, but the spring constants are different.

Let x_1 be the displacement of the first mass from its equilibrium and x_2 be the displacement of the second mass from its equilibrium. To work out Newton's laws, we first want to know the force on x_1 when it is moved from its equilibrium while holding x_2 fixed. This is

$$F_{\text{on 1 from moving 1}} = F = -kx_1 - \kappa x_1 \tag{2}$$

The signs are both chosen so that they oppose the motion of the mass. There is also a force on x_1 if we move x_2 holding x_1 fixed. This force is

$$F_{\text{on 1 from moving 2}} = \kappa x_2 \tag{3}$$

To check the sign, note that if x_2 is increased, it pulls x_1 to the right. There is no contribution to this force from the spring between the second mass and the wall, since we are moving the mass by hand and just asking how it affects the first mass. Thus

$$m\,\ddot{x}_1 = -(k+\kappa)x_1 + \kappa x_2 \tag{4}$$

similarly,

$$m\,\ddot{x}_2 = -(k+\kappa)x_2 + \kappa x_1 \tag{5}$$

One way to solve these equations is to note that if we add them, we get

$$m(\ddot{x}_1 + \ddot{x}_2) = -k(x_1 + x_2) \tag{6}$$

This is just $m\,\ddot{y} = -ky$ for $y = x_1 + x_2$, so the solutions are sines and cosines, or cosine and a phase:

$$x_1 + x_2 = A_s \cos(\omega_s t + \phi_s), \quad \omega_s = \sqrt{\frac{k}{m}} \tag{7}$$

Another way solve them is taking the difference

$$m(\ddot{x}_1 - \ddot{x}_2) = (-k - 2\kappa)(x_1 - x_2) \quad \Rightarrow \quad x_1 - x_2 = A_f \cos(\omega_f t + \phi_f), \quad \omega_f = \sqrt{\frac{k + 2\kappa}{m}} \tag{8}$$

We write ω_s for ω_{slow} and ω_f for ω_{fast}, since $\omega_f > \omega_s$. Thus we have found two solutions each of which oscillate with fixed frequency. These are the **normal modes** for this system. A general solution is a linear combination of these two solutions. Explicitly, we have:

$$x_1 = \frac{1}{2}[(x_1 + x_2) + (x_1 - x_2)] = \frac{1}{2}[A_s \cos(\omega_s t + \phi_s) + A_f \cos(\omega_f t + \phi_f)] \tag{9}$$

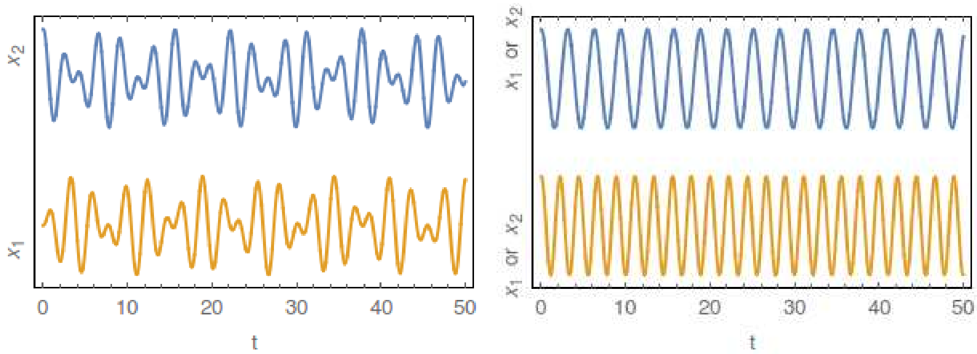
$$x_2 = \frac{1}{2}[(x_1 + x_2) - (x_1 - x_2)] = \frac{1}{2}[A_s \cos(\omega_s t + \phi_s) - A_f \cos(\omega_f t + \phi_f)] \tag{10}$$

If we can excite the masses so that $A_f = 0$ then the masses will both oscillate at the frequency ω_s. In practice, we can do this by pulling the masses to the right by the same amount, so that $x_1(0) = x_2(0)$ which implies $A_f = 0$. The solution is then $x_1 = x_2$ and both oscillate at the frequency A_s for all time. This is the **symmetric oscillation mode.** Since $x_1 = x_2$ at all times, both masses move right together, then move left together.

If we excite the masses in such a way that $A_s = 0$ then $x_1 = -x_2$ and both oscillate at frequency ω_f. We can set this up by pulling the masses in opposite directions. In this mode, when one mass is right of equilibrium, the other is left, and vice versa. So this is an **antisymmetric mode**.

2 Beats

You should try playing with the coupled oscillator solutions in the Mathematica notebook oscillators.nb. Try varying κ and k to see how the solution changes. For example, say $m = 1$, $\kappa = 2$ and $k = 4$. Then $\omega_s = 2$ and $\omega_f = 2\sqrt{2}$, Here are the solutions:

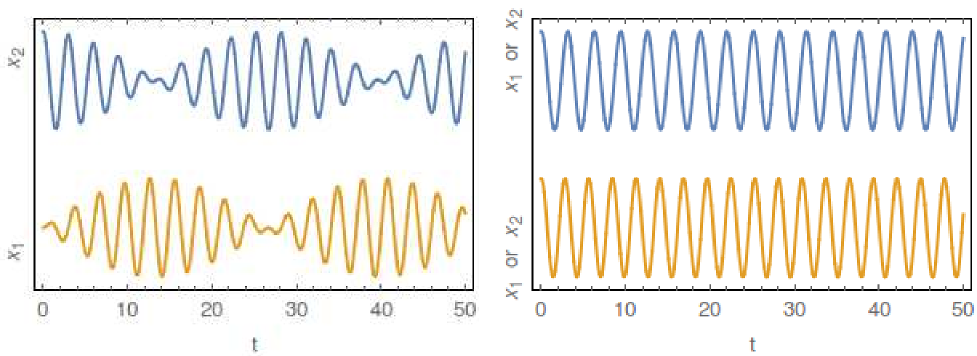


Behavior starting from $x_1 = 1, x_0 = 0$ Normal mode behavior

Figure 1. Left shows the motion of masses $m = 1, \kappa = 2$ and $k = 4$ starting with $x_1 = 1$ and $x_2 = 0$. Right shows the normal modes, with $x_1 = x_2 = 1$ (top) and $x_1 = 1, x_2 = -1$ (bottom).

If you look closely at the left plot, you can make out two distinct frequencies: the normal mode frequencies, as shown on the right.

Now take $\kappa = 0.5$ and $k = 4$. Then $\omega_s = 2$ and $\omega_f = 2.2$. In this case



Behavior starting from $x_1 = 1, x_0 = 0$ Normal mode behavior

Figure 2. Motion of masses and normal modes for $k = 0.5$ and $\kappa = 4$

Now we can definitely see two distinct frequencies in the positions of the two masses. Are these the two frequencies ω_s and ω_f? Comparing to the normal mode plots, it is clear they are not. One is much slower. However, we do note that $\omega_s \approx \omega_f$. What we are seeing here is the emergence of **beats**. Beats occur when two normal mode frequencies get close.

Beats can be understood from the simple trigonometric relation

$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2\cos\left(\frac{\omega_1 + \omega_2}{2} t\right)\cos\left(\frac{\omega_1 - \omega_2}{2} t\right) \tag{11}$$

When you excite two frequencies ω_1 and ω_2 at the same time, the solution to the equations of motion is the sum of the separate oscillating solutions (by linearity!). Eq. (11) shows that this sum can also be written as the *product* of two cosines. In particular, if $\omega_1 \approx \omega_2$ then

$$\omega = \frac{\omega_1 + \omega_2}{2} \approx \omega_1 \approx \omega_2 \quad \varepsilon = \frac{\omega_1 - \omega_2}{2} \ll \omega_1, \omega_2 \tag{12}$$

So the sum looks like an oscillation whose frequency ω is the *average* of the two normal mode frequencies modulated by an oscillation with frequency ε given by half the difference in the frequencies.

Beats are important because they can generate frequencies well below the normal mode frequencies. For example, suppose you have two strings which are not quite in tune. Say they are supposed to both be the note $A4$ at 440 Hz, but one is actually $\nu_1 = 442\,\mathrm{Hz}$ and the other is $\nu_2 = 339\,\mathrm{Hz}$. If you pluck both strings together you will hear the average frequency $\Omega = 440.5\,\mathrm{Hz}$, but also there will be an oscillation at $\varepsilon = \frac{1}{2}(442 - 339)\mathrm{Hz} = 1.5\,\mathrm{Hz}$. This oscillation is the enveloping curve over the high frequency (440.5 Hz) oscillations

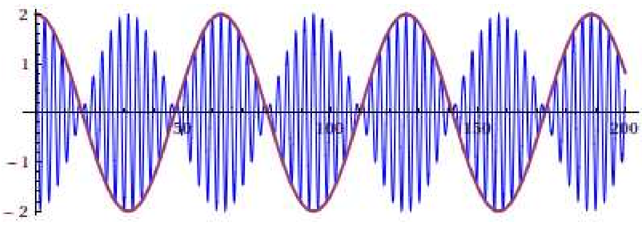


Figure 3. The red curve is $\cos\left(2\pi \frac{\nu_1 - \nu_2}{2} t\right)$. When hearing beats, the observed frequency is the frequency of the extrema $\nu_{\text{beat}} = \nu_1 - \nu_2$ which is twice the frequency of this curve .

As you can see from the figure, due to the high frequency oscillations, there are peaks in the amplitude twice as often as peaks in $\cos\left(2\pi \frac{\nu_1 - \nu_2}{2} t\right)$. Thus what we hear are beats at the **beat frequency**

$$\nu_{\text{beat}} = |\nu_1 - \nu_2| \tag{13}$$

We use an absolute value since we want a frequency to be positive (it's the same frequency whether $\nu_1 > \nu_2$ or $\nu_2 > \nu_1$). Note that there is no factor of 2 in the conventional definition of ν_{beat}, since we only ever hear the modulus of the oscillation not the phase.

Thus with $\nu_f = 442\,\mathrm{Hz}$ and $\nu_s = 339\,\mathrm{Hz}$ the beat frequency is $\nu_{\text{beat}} = 3\,\mathrm{Hz}$. Thus you hear something happening 3 times a second. This is a regular beating in off-tune notes which is audible by ear. In fact, it is a useful trick for tuning – change one string until the beating disappears. Then the strings are in tune. We will see numerous examples of beats as the course progresses.

3 Two masses with matrices

We solved the two coupled mass problem by looking at the equations and noting that their sum and difference would be independent solutions. For more complicated systems (more masses, different couplings) we should not expect to be able to guess the answer in this way. Can you guess the solution if the two oscillators have different masses?

To develop a more systematic procedure, suppose we have lots of masses with lots of different springs connected in a complicated way. Then the equations of motion are

$$m_1\ddot{x_1} = k_{11}\,x_1 + k_{12}\,x_2 + \cdots + k_{1n}x_n \tag{14}$$

$$\cdots \tag{15}$$

$$m_n\ddot{x_n} = k_{n1}\,x_1 + k_{n2}\,x_2 + \cdots + k_{\text{nn}}x_n \tag{16}$$

where k_{ij} are constants, representing the strength of the spring between masses i and j. Note that all of these equations are linear. What are the solutions in this general case? This is an algebra problem involving linear equations. Hence we should be able to solve it with **linear algebra**.

To connect to linear algebra, let's return to our two mass system. Since the equations of motion are linear, we expect them to be solved by exponentials $x_1 = c_1 e^{i\omega t}$ and $x_2 = c_2 e^{i\omega t}$ for some ω, c_1 and c_2. As with the driven oscillator from the last lecture, we are using complex solutions to make the math simpler, then we can always take the real part at the end. Plugging in these guesses, Eqs. (4) and (5) become

$$-m_1\omega^2 c_1 = -(k+\kappa)c_1 + \kappa c_2 \tag{17}$$

$$-m_2\omega^2 c_2 = -(k+\kappa)c_2 + \kappa c_1 \tag{18}$$

We have let the masses be different for generality.

Next, we will write these equations in matrix form. To do so, we define a vector $\vec{c}$ as

$$\vec{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{19}$$

Then the equations of motion become

$$M\cdot\vec{c} = \begin{pmatrix} \frac{-k-\kappa}{m_1} & \frac{\kappa}{m_1} \\ \frac{\kappa}{m_2} & \frac{-k-\kappa}{m_2} \end{pmatrix}\cdot\vec{c} = -\omega^2\vec{c} \tag{20}$$

where M is defined by this equation.

You might recognize this as an eigenvalue equation. An $n\times n$ matrix A has n **eigenvalues** λ_i and n associated **eigenvectors** $\vec{v}_i$ which satisfy

$$A\cdot\vec{v}_i = \lambda_i\vec{v}_i \tag{21}$$

The eigenvalues don't all have to be different. Note that the left hand side is a matrix multiplying a vector while the right-hand side is just a number multiplying a vector. So studying eigenvalues and eigenvectors lets us turn matrices into numbers! Eigenvalues and eigenvectors are *the* fundamental mathematical concept of quantum mechanics. I cannot emphasize enough how important it is to master them.

Let's recall how to solve an eigenvalue equation. The trick is to write it first as

$$(A-\lambda\mathbb{1})\vec{v} = 0 \tag{22}$$

where $\mathbb{1}$ is the $n\times n$ identity matrix. For $n=2$, $\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. For most values of λ, the matrix $(A-\lambda\mathbb{1})$ has an inverse. Multiplying both sides of Eq. (22) by that inverse, we find $\vec{v}=0$. This is the trivial solution (it obviously satisfies Eq. (21) for any A). The nontrivial solutions consequently must correspond to values of λ for which $(A-\lambda\mathbb{1})$does **not** have an inverse. When does a matrix not have an inverse? A result from linear algebra is that a matrix is not invertible if and only if its determinant is zero. Thus the equation $\det(A-\lambda\mathbb{1})=0$ is an algebraic equation for λ whose solutions are the eigenvalues λ_i.

It is useful to know that determinant of a 2×2 matrix is

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{23}$$

You should have this memorized. For a 3×3 matrix, the determinant is:

$$\det\begin{pmatrix} a & b & c\\ d & e & f\\ g & h & i \end{pmatrix} = a(ei-fh)-b(di-fg)+c(dh-eg) \tag{24}$$

You should know how to compute this, but don't need to memorize the formula. Beyond 3×3, you probably want to take determinants with Mathematica rather than by hand.

So, returning to Eq. (20), the eigenvalues $-\omega^2$ must satisfy

$$0=\det(M+\omega^2\mathbb{1})=\det\begin{pmatrix} \frac{-k-\kappa}{m_1}+\omega^2 & \frac{\kappa}{m_1}\\ \frac{\kappa}{m_2} & \frac{-k-\kappa}{m_2}+\omega^2 \end{pmatrix} \tag{25}$$

$$=\left(\frac{-k-\kappa}{m_1}+\omega^2\right)\left(\frac{-k-\kappa}{m_2}+\omega^2\right)-\frac{\kappa^2}{m_1m_2} \tag{26}$$

This is a quadratic equation for ω^2, with two roots: the two eigenvalues.

Let's set $m_1=m_2=m$ now to check that we reproduce our old result. Multiplying Eq. (26) by m^2, it reduces to

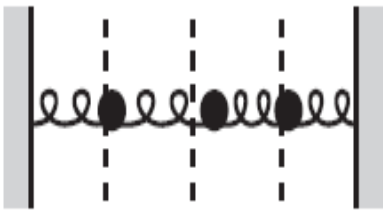
$$(k+\kappa-m\omega^2)^2=\kappa^2 \tag{27}$$

Thus $k+\kappa-m\omega^2=\pm\kappa$. Or in other words

$$\omega=\omega_s=\sqrt{\frac{k}{m}},\quad \omega=\omega_f=\sqrt{\frac{k+2\kappa}{m}} \tag{28}$$

These are the two normal mode frequencies we found above. Note that we didn't have to take the real part of the solution to find the normal mode frequencies. We only need to take the real part to find the solutions $x(t)$.

Now let's try three masses. We can couple them all together and to the walls in any which way



(29)

The equations of motion for this system will be of the form

$$m_1\ddot{x}_1=k_{11}x_1+k_{12}x_2+k_{13}x_3 \tag{30}$$

$$m_2\ddot{x}_3=k_{21}x_1+k_{22}x_2+k_{23}x_3 \tag{31}$$

$$m_3\ddot{x}_3=k_{31}x_1+k_{32}x_2+k_{33}x_3 \tag{32}$$

Some of these k_{ij} are probably zero, but we don't care. Writing $x_1=c_1e^{i\omega t}$, $x_2=c_2e^{i\omega t}$ and $x_3=c_3e^{i\omega t}$, these equations become algebraic:

$$-\omega^2c_1=\frac{k_{11}}{m_1}c_1+\frac{k_{12}}{m_1}c_2+\frac{k_{13}}{m_1}c_3 \tag{33}$$

$$-\omega^2c_2=\frac{k_{21}}{m_2}c_1+\frac{k_{22}}{m_2}c_2+\frac{k_{23}}{m_2}c_3 \tag{34}$$

$$-\omega^2c_3=\frac{k_{31}}{m_3}c_1+\frac{k_{32}}{m_3}c_2+\frac{k_{33}}{m_3}c_3 \tag{35}$$

In other words,

$$(M+\omega^2\mathbb{1})\vec{x}=0 \tag{36}$$

with M the matrix whose entries are $M_{ij} = \frac{k_{ij}}{m_i}$. So to find the normal mode frequencies ω, we need to solve $\det(M + \omega^2 \mathbb{1}) = 0$. For a 3×3 matrix, there will be 3 eigenvalues and hence three normal-mode frequencies.

The Noether theorem

We already saw that if q is a cyclic variable, the associated conjugate momentum is conserved,

$$\frac{\partial \mathcal{L}}{\partial q} = 0 \quad \Rightarrow \quad p \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}} = \text{const} \; . \tag{1}$$

This is the simplest incarnation of Noether's theorem, which states that whenever we have a *continuous symmetry* of Lagrangian, there is an associated *conservation law*. By "symmetry" we mean any transformation of the generalized coordinates q, of the associated velocities $\dot{q}$, and possibly of the time variable t, that leaves the value of the Lagrangian unaffected. By "continuous symmetry" we mean a symmetry with a continuous constant parameter, typically infinitesimal, say ϵ, that we can dial, and that measures how far from the identity the transformation is bringing us. In a sense ϵ measures the "size" of the transformation.

In the case of the cyclic coordinate discussed above, the corresponding symmetry is simply

$$q(t) \to q(t) + \epsilon \; , \qquad \dot{q}(t) \to \dot{q}(t) \; , \qquad t \to t \; , \tag{2}$$

that is, an infinitesimal shift of the cyclic coordinate. Indeed, if we perform these replacements in the Lagrangian, at first order in ϵ the Lagrangian changes by

$$\delta \mathcal{L} \equiv \mathcal{L}(q + \epsilon, \dot{q}; t) - \mathcal{L}(q, \dot{q}; t) \simeq \frac{\partial \mathcal{L}}{\partial q} \epsilon \; , \tag{3}$$

which vanishes if an only if q is cyclic.

Theorem: Consider a Lagrangian system with n degrees of freedom $q_1, \ldots, q_n$. *If* for certain functions $\gamma_\alpha(t)$ and for constant infinitesimal ϵ the transformation

$$q_\alpha(t) \to q_\alpha(t) + \epsilon \gamma_\alpha(t) \; , \qquad \dot{q}_\alpha(t) \to \dot{q}_\alpha(t) + \epsilon \dot{\gamma}_\alpha(t) \; , \qquad t \to t \; , \tag{4}$$

is a symmetry, i.e. if it leaves the Lagrangian unaffected, *then* the quantity

$$\sum_{\alpha=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \gamma_\alpha \tag{5}$$

is a constant of motion, i.e. it is conserved. Notice that in (30) we are not transforming the time variable. We will treat the case of time transformations, in particular of time translations, separately below.

Proof: By definition of symmetry, the change in the Lagrangian upon the replacements (30) must vanish

$$\delta \mathcal{L} \equiv \mathcal{L}(q_\alpha + \epsilon \gamma_\alpha, \dot{q}_\alpha + \epsilon \dot{\gamma}_\alpha; t) - \mathcal{L}(q_\alpha, \dot{q}_\alpha; t) = 0 \; . \tag{6}$$

At first order in ϵ, this equation becomes

$$\sum_\alpha \left[\frac{\partial \mathcal{L}}{\partial q_\alpha} \epsilon \gamma_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \epsilon \dot{\gamma}_\alpha \right] = 0 \; . \tag{7}$$

We can rewrite the first term by using the equations of motion:

$$\frac{\partial \mathcal{L}}{\partial q_\alpha} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\ , \qquad \forall \alpha\ . \tag{8}$$

We are left with

$$\sum_\alpha \left[\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\epsilon\gamma_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\epsilon\dot{\gamma}_\alpha\right] = 0\ . \tag{9}$$

The l.h.s. we can rewrite as a total time derivative

$$\epsilon\frac{d}{dt}\Big(\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\gamma_\alpha\Big) = 0 \tag{10}$$

This implies that the quantity (5) is conserved:

$$\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\gamma_\alpha = \text{const}\ . \tag{11}$$

□

Examples

Euclidean translations. Consider N point particles, interacting via a potential. The Lagrangian is

$$\mathcal{L} = \sum_{a=1}^N \tfrac{1}{2}m_a\dot{\vec{r}}_a^{\,2} - V(\vec{r}_1, \ldots, \vec{r}_N)\ . \tag{12}$$

If the potential only depends on the relative positions $\vec{r}_a - \vec{r}_b$, and not on the absolute ones (i.e., if there are no *external* forces),

$$V = V(\vec{r}_1 - \vec{r}_2, \ldots, \vec{r}_1 - \vec{r}_N, \vec{r}_2 - \vec{r}_3, \ldots)\ , \tag{13}$$

then *overall translations* of the system are a symmetry. An infinitesimal translation of length ϵ in an arbitrary direction $\hat{n}$ takes the form

$$\vec{r}_a \to \vec{r}_a + \epsilon\hat{n}\ , \qquad \dot{\vec{r}}_a \to \dot{\vec{r}}_a\ , \qquad t \to t \qquad \forall a\ . \tag{14}$$

The corresponding conserved quantity is thus

$$\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\gamma_\alpha = \sum_{a=1}^N\sum_{i=1}^3 \frac{\partial \mathcal{L}}{\partial \dot{r}_a^i}\hat{n}^i = \text{const} \tag{15}$$

(the role of the γ_α's is played by the cartesian components of $\hat{n}$.) The direction $\hat{n}$ is the same for all particles and we can thus pull it out of the sum over a:

$$\sum_{i=1}^{3} \hat{n}^i \sum_{a=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{r}_a^i} = \text{const} \tag{16}$$

From (12) we have

$$\frac{\partial \mathcal{L}}{\partial \dot{r}_a^i} = m_a \dot{r}_a^i \tag{17}$$

so that

$$P^i \equiv \sum_{a=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{r}_a^i} \tag{18}$$

is nothing but the i-th component of the total momentum of the system. Our conservation law thus takes the form

$$\hat{n} \cdot \vec{P} = \text{const} \ . \tag{19}$$

Since $\hat{n}$ is an arbitrary direction, the whole vector $\vec{P}$ should be constant

$$\vec{P} = \vec{\text{const}} \ . \tag{20}$$

We therefore see that the conservation of the total momentum in the absence of external forces is a direct consequence of the invariance of the Lagrangian under spacial translations.

Euclidean rotations. If we make the further assumption that the potential only depends of the mutual *distances* $|\vec{r}_a - \vec{r}_b|$ between the particles, and not on the orientation of the relative position vectors $\vec{r}_a - \vec{r}_b$,

$$V = V(|\vec{r}_1 - \vec{r}_2|, \ldots, |\vec{r}_1 - \vec{r}_N|, |\vec{r}_2 - \vec{r}_3|, \ldots) \ , \tag{21}$$

then the Lagrangian is also invariant under *overall rotations* of the system, because the potential is, and the kinetic energy is also since it only involves the scalar quantities $\dot{\vec{r}}_a^{\,2}$. An infinitesimal rotation of angle ϵ about an arbitrary axis $\hat{n}$ takes the form

$$\vec{r}_a \to \vec{r}_a + \epsilon\, \hat{n} \times \vec{r}_a \ , \qquad \dot{\vec{r}}_a \to \dot{\vec{r}}_a + \epsilon\, \hat{n} \times \dot{\vec{r}}_a \ , \qquad t \to t \qquad \forall a \ . \tag{22}$$

The corresponding conserved quantity is

$$\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \gamma_\alpha = \sum_{a=1}^{N} \sum_{i=1}^{3} \frac{\partial \mathcal{L}}{\partial \dot{r}_a^i} (\hat{n} \times \vec{r}_a)^i = \text{const} \tag{23}$$

(the role of the γ_α's is played by the cartesian components of $(\hat{n} \times \vec{r}_a)$.) Using (17) we get

$$\sum_{a=1}^{N} m_a \dot{\vec{r}}_a \cdot \big(\hat{n} \times \vec{r}_a\big) = \text{const} \ . \tag{24}$$

For any three vectors $\vec{A}, \vec{B}, \vec{C}$, the following identity holds:

$$\vec{A} \cdot (\vec{B} \times \vec{C}) = \vec{B} \cdot (\vec{C} \times \vec{A}) \tag{25}$$

We can thus rewrite our conservation law as

$$\sum_{a=1}^{N} \hat{n} \cdot \left(\vec{r}_a \times m_a \dot{\vec{r}}_a\right) = \text{const} \; . \tag{26}$$

The direction $\hat{n}$ is the same for all particles. We can then pull it out of the sum, and we recognize in the remainder the *total angular momentum* of the system

$$\vec{L} \equiv \sum_{a=1}^{N} \vec{r}_a \times m_a \dot{\vec{r}}_a \; . \tag{27}$$

Our conservation law becomes

$$\hat{n} \cdot \vec{L} = \text{const} \; , \tag{28}$$

or, since the direction $\hat{n}$ is arbitrary,

$$\vec{L} = \vec{\text{const}} \; . \tag{29}$$

The conservation of the total angular momentum is a direct consequence of the invariance of the Lagrangian under overall rotations of the system.

Time translations. This case has to be treated separately because our simplified formulation of the general theorem does not cover it. By "time-translation" we mean the transformation

$$q_\alpha(t) \to q_\alpha(t) \; , \qquad \dot{q}_\alpha(t) \to \dot{q}_\alpha(t) \; , \qquad t \to t + \epsilon \; . \tag{30}$$

That is, we do nothing to the coordinates and to the velocities, but we shift time by an infinitesimal constant. This is a symmetry if and only if the Lagrangian does not depend *explicitly* on time. Indeed, the variation of the Lagrnagian under the above transformation would be

$$\delta\mathcal{L} \equiv \mathcal{L}(q, \dot{q}; t + \epsilon) - \mathcal{L}(q, \dot{q}; t) \simeq \frac{\partial \mathcal{L}}{\partial t} \epsilon \; , \tag{31}$$

which vanishes if and only if the partial time-derivative of the Lagrangian vanishes. Recall that the *total* time-derivative of the Lagrangian does not vanish in general, because on any given solution the value of Lagrangian depends on time also through the time dependence of the q's and of the $\dot{q}$'s. One has

$$\frac{d}{dt}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial t} + \sum_\alpha \left[\frac{\partial \mathcal{L}}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\ddot{q}_\alpha\right] \tag{32}$$

We can rewrite the first term inside the brackets via the eom (8). We get

$$\frac{d}{dt}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial t} + \sum_\alpha \left[\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\dot{q}_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\ddot{q}_\alpha\right] \tag{33}$$

We then notice that the two terms inside the brackets combine to give a total time derivative:

$$\frac{d}{dt}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt}\sum_{\alpha} \frac{\partial \mathcal{L}}{\partial \dot{q}_{\alpha}}\dot{q}_{\alpha} \tag{34}$$

This equation is more conveniently rewritten as

$$\frac{d}{dt}H = -\frac{\partial \mathcal{L}}{\partial t} \ , \tag{35}$$

where we defined H, the *Hamiltonian* of the system, as

$$H \equiv \sum_{\alpha} \frac{\partial \mathcal{L}}{\partial \dot{q}_{\alpha}}\dot{q}_{\alpha} - \mathcal{L} \ . \tag{36}$$

In summary, equation (35) is always valid, but *if* the Lagrangian is invariant under time-translations, that is if it does not depend explicitly on time, then the Hamiltonian of the system is conserved

$$H = \text{const} \ . \tag{37}$$

In most physically relevant cases the value of the Hamiltonian is the *total energy*. We thus discovered that the conservation of energy is a direct consequence of the invariance of the Lagrangian under time translations. Under stable conditions, if you perform a lab experiment today or tomorrow you expect to get the same results. This fact alone implies that energy is conserved.